AJS

CONTENTS

New Directions in the Sociology of Medicine

Theoretical Implications of Health Behavior

Book Reviews

K-P-12   Aw 35

ERRATA

AJS Erratum

In the July 1991 issue (*AJS* 97 [1]:237–38), Linda Lobao's name was misspelled in a review of her book, *Locality and Inequality*. We regret the error.

Authors' Errata

John Hagan and Alberto Palloni wish to clarify two equations that appeared in their article, "The Social Reproduction of a Criminal Class in Working-Class London, circa 1950–1980" (*AJS* 96[2]:265–99). In equation (A2), a superscript "n" should replace the superscript "m"; in equation (A4), "$+ U$" should appear between the closing parenthesis and the closing bracket to the first set of fences. The two equations are thus:

$$P(n{:}Z_j) = [\exp (B \times Z_j) \times k_0]^n \times \exp[-k_0 \times \exp(B \times Z_j)] \times (n!)^{-1}, \tag{A2}$$

and

$$P(n) = [U/(k_0 \times H_j) + U]^W \times \{(k_0 \times H_j)/[(k_0 \times H_j) + U]\}^n \times W \tag{A4}$$
$$\times (W - 1)\ldots(W + n - 1) \times (n!)^{-1}.$$

In both cases, the problem arose from typographical errors in the manuscript.

IN THIS ISSUE

DONALD W. LIGHT is a professor at both the University of Medicine and Dentistry of New Jersey and the graduate faculty of Rutgers University. He is completing a historical analysis of the relations between the profession, state, and economy in the American health care system during the 20th century. He led a team that authored a similar study of the German system entitled *Political Values and Health Care: The German Experience* (MIT Press, 1986) and is gathering materials for a study of the British system.

MARY RUGGIE is professor of sociology and public affairs at Columbia University. Her most recent publication is "Retrenchment or Realignment? U.S. Mental Health Policy and DRGs" (*Journal of Health Politics, Policy and Law* 15 [Spring 1990]). She is working on a comparative study of health policy in the United States, Britain, and Canada.

LEE CLARKE is assistant professor of sociology at Rutgers and author of *Acceptable Risk? Making Decisions in a Toxic Environment* (Berkeley and Los Angeles: University of California Press, 1989), an article entitled "Oil Spill Fantasies," which appeared in the November 1990 issue of the *Atlantic Monthly*, and a number of articles on organizations and technological risk. His book on organizational responses to extreme uncertainty, tentatively titled *Controlling the Uncontrollable,* will be published by the University of Chicago Press.

CARROLL L. ESTES is professor of sociology and chair of the Department of School and Behavioral Sciences at the University of California, San Francisco. She is also director of the Institute for Health and Aging at the UCSF School of Nursing. Among her many authored or coauthored books is the forthcoming *Health Care Access and Post-hospital Care for the Elderly.*

ALLEN W. IMERSHEIN is director of the Institute for Health and Human Services Research and holds appointments in the Department of Sociology and the School of Public Administration at Florida State University. The article in this issue is his second effort in an ongoing study of health elite restructuring. His other research includes studies on organizational reform in human services with a focus on workers' compensation systems.

PHILIP C. ROND III is senior research associate at the Institute for Health and Human Services Research, Florida State University, where he has engaged in a variety of state health policy studies. He also has extensive experience as a health planner, health program administrator, and health policy consultant. His current research interests relate to the politics of state health policy formation and implementation.

MARY P. MATHIS received her Ph.D. in sociology from Florida State University in 1990 and is currently doing postdoctoral work at Emory University and the Centers for Disease Control. Her interests are in child health and HIV/AIDS. She has recently worked as a research associate at the Institute for Health and Human Services Research at Florida State University.

SYDNEY A. HALPERN is associate professor of sociology at the University of Illinois at Chicago. Much of her work concerns social change in medical organizations and professions. She is currently studying the division of labor among health care occupations.

MARTHA S. GERRITY, M.D., is a clinical instructor in medicine and a National Service Award (Agency for Health Care Policy and Research) fellow at the University of North Carolina at Chapel Hill. Her research interests include the influences of uncertainty and medical education on physicians' attitudes and clinical practice and measurement methodology in health care.

JO ANNE L. EARP is associate professor in the Department of Health Behavior and Health Education in the School of Public Health at the University of North Carolina at Chapel Hill. Her research interests center on the sociology of health care delivery and utilization of preventive services, with a special emphasis on women and underserved populations. She also has been studying the effects of counseling, testing, and community interventions on people at risk of AIDS and other sexually transmitted diseases.

ROBERT F. DeVELLIS is assistant director of the Rehabilitation Program in the School of Medicine, research associate professor in the Department of Health Behavior and Health Education (School of Public Health) and the Department of Psychology (School of Arts and Sciences), and a research associate of the Cecil B. Sheps Center for Health Services Research, all at the University of North Carolina at Chapel Hill. Much of his research involves assessing social psychological processes and examining their relation to health and illness.

ELLEN L. IDLER is associate professor of sociology and a member of the Institute for Health, Health Care Policy, and Aging Research at Rutgers University. She studies the effect of cultural and social life on morbidity, mortality, and subjective assessments of health. Her current research with a disabled population examines discrepancies between physical health and perceptions of it.

STANISLAV V. KASL is professor of epidemiology and director of graduate studies in the Department of Epidemiology and Public Health, Yale University School of Medicine. His primary research interest has been in social and psychological risk factors for disease and adverse changes in physical health.

THOMAS A. LAVEIST is assistant professor of health policy and management and Brookdale National Fellow at the School of Hygiene and Public Health at Johns Hopkins University. In 1989, he received the outstanding dissertation award given by the Medical Sociology Section of the American Sociological Association. He is currently conducting research on the social and behavioral determinants of longevity among African-Americans.

BERNICE A. PESCOSOLIDO is associate professor of sociology at Indiana University. Under a research scientist development award from the National Institute of Mental Health and research grants from NIMH and the National Science Foundation, she is pursuing her work on the role of social networks in decision-making processes, particularly those involving health, illness, and healing.

Information for Contributors

Editorial Procedures: All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. All identifying references and footnotes should appear on a separate page. A legible, carefully prepared manuscript will facilitate the work of readers. Manuscripts are accepted for publication subject to nonsubstantive editing with the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his article without charge in any book of which he is the author or editor after it has appeared in the *Journal*.

Manuscript Acceptance Policy: While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the journal has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the journal in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

Preparation of Copy

1 Type *all* copy—including indented matter, footnotes, tables, and references—*double-spaced,* allowing generous margins at top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Avoid proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, including acknowledgment footnote (which should be "1"). *Provide a word count on the cover page of the manuscript.*

2 Type each table on a separate page. Refer to each table in numerical order in the text. Prepare tables without vertical lines. Tables should not contain more than 20 two-digit columns or the equivalent.

3 Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Draw figures on white paper with india ink. Original or glossy print of professionally drawn figure will be required if manuscript is accepted

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6 *Five copies* of the manuscript and abstract, on white paper, must be submitted. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, The University of Chicago, Chicago, Illinois 60637. *Manuscripts are not returned to authors* Enclose a self-addressed stamped postcard, containing the manuscript title, to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or money order for $15.00, payable to The University of Chicago Press in U.S. currency or its equivalent by money order or bank draft. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

Format of References in Text: All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text· (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61–65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (*a, b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons.

Format of References: List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "References." Examples follow:

Davis, James Allan. 1978. *General Social Survey, 1972–1978: Cumulative Data* (MRDF). NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

Davis, K. 1963*a*. "Social Demography." Pp. 124–37 in *The Behavioral Sciences Today,* edited by Bernard Berelson. New York. Basic.

————. 1963*b*. "The Theory of Change and Response in Modern Demographic History." *Population Index* 29.345–66.

Goode, W. J. 1967. "The Protection of the Inept." *American Sociological Review* 32.5–19.

Moore, Wilbert E., and Arnold S. Feldman. 1960. *Labor Commitment and Social Change in Developing Areas.* New York: Social Science Research Council.

Sanford, Nevitt, ed. 1962. *The American College.* New York: Wiley.

U.S. Bureau of the Census. 1970. *Census of Population and Housing 1970, Summary Statistic File 4H: U.S.* (MRDF). DUALabs ed. Washington, D.C.. Bureau of the Census. Distributed by DUALabs, Rosslyn, Va.

Weber, Max. (1921) 1968. "Society's Problems." Pp. 12–16 in *Economy and Society,* edited by Guenther Roth and Claus Wittich. New York: Bedminister.

(Rev. 7/91)

Introduction: Strengthening Ties between Specialties and the Discipline[1]

Donald W. Light, Contributing Editor
*University of Medicine and Dentistry of New Jersey
and Rutgers University*

ORIGINS

In organizing a session on the sociology of American medicine for the discipline's annual meeting in 1987, I was surprised to discover many papers that took fresh approaches to old topics, recast theories that had pervaded the field for 20 years, or mapped entirely new territory. The sociology of medicine, it appeared, was undergoing one of those generative periods when fresh gales blow through a specialty, uprooting old plants and bringing in new seed. For too long, it seemed to me, this large and fertile field had been working off old plantings, and further inquiry convinced me that younger scholars and a few established ones were doing more original work than one normally sees. This warranted special attention: a new body of work was being produced that drew upon and contributed to the discipline as a whole.

Many of the early submissions for this issue did not pass muster, but for encouraging reasons. They broke new ground, extended the domain of the field, or brought a fresh perspective; but reviewers judged them too speculative in one sense of that word or another. This was a sad outcome and personally painful to one who had worked for hours and weeks with several of the authors. The experience implies that the attempt by a major journal to be definitive impedes its desire and obligation to be original. Perhaps a new section should be created entitled "New Ideas," where original observations or concepts could be presented briefly. This would both inform colleagues of new ideas so that their generative power can begin to influence the work of others, and it would establish the authorship of original thinking or observation well before it can be thoroughly researched. This latter advantage would additionally allow the discipline to benefit from the original thinking among a large

cadre of colleagues whose circumstances make it unlikely that they would *ever* have the resources to turn a great idea into a full-blown article.

Given that even the rejected papers attested to the freshness of new work in the sociology of medicine, I continued to work with authors who submitted promising drafts. The result is a collection from largely new voices. Only one among the first authors in this group had secured tenure when the papers were submitted. This is, after all, what makes academic work exciting: new ideas, new findings, new authors.

THE DANGERS OF SPECIALTY ISOLATION

Before turning to the articles themselves, it is worth reflecting on the dangers that sociology faces from specialization. Randall Collins (1986) has written with insight and eloquence about the social dynamics of specialization. As membership in the American Sociological Association (ASA) more than doubled from 1950 to 1960 and then more than doubled again by 1970 (before it flattened out and started a slow decline from its peak), large numbers of faculty used specialization as the strategy for recognition and originality. "The deepest part of our intellectual malaise, then, is due to crises in the making of intellectual careers today. . . . We have become congeries of outsiders to each other" (1986, p. 1340). Specialists burrow in, work and rework the veins of ore in their tunnels, and, if they look up from their labors, cannot see much. This self-isolation naturally arises, Collins writes, as a given group of specialists refine their methods and ideas. They become self-referent and no longer feel the need to draw upon or contribute to the larger academic study of social life. In so doing, they also become of less and less interest to other sociologists. Both the larger discipline and the specialty lose what each might give the other. This problem is particularly acute in specialties of application, like the sociology of education, or of medicine, or of sports, that need the power of larger theories and methods to inform their work.

THE CASE OF MEDICAL SOCIOLOGY

Medical sociology, though the largest section of the ASA and fully institutionalized (Bloom 1986), runs the danger of self-isolation, which makes it vulnerable to being taken over by competing disciplines. Most medical sociology concerns the epidemiology and social correlates of morbidity, mortality, illness behavior, stress, risk factors, and other social psychological aspects that are ripe for takeover by psychologists, epidemiologists, and/or clinicians. The predominance of this work stems from a long effort by medical sociologists to be accepted by the medical community and medically oriented funders (Freeman and Reeder 1957); ironi-

cally, acceptance has brought with it a loss of sociological distinctiveness and power. This work is well funded, and indeed some sociology departments define their medical sociology slot in terms of this work and the grants it can bring to the department.

The sociology of illness can be mistaken for the entire field, an assumption that is reflected in Uta Gerhardt's (1989) excellent book, *Ideas about Illness,* which is described as an "intellectual and political history of medical sociology." The title and the book leave out a powerful and rich body of research on the organizational and institutional aspects of services, training, and research. Even the treatment of fascism takes on a psychiatric cast and focuses on the question of whether mentally healthy people can support monstrous political institutions.

One is reminded of C. Wright Mills's (1943) classic article about the dangers of focusing on process and deviating from norms and situational adjustments so that institutional structures and the distributions of power that shape them go unexamined. Fifteen years later, Ralf Dahrendorf (1958) forcefully wrote in these pages about "objective disengagement," sociology without history, and the general irrelevance of sociology that did not address the central role of conflicts and constraints in social life.

These dangers in the study of illness behavior, stress, and epidemiology can be avoided, as illustrated by Waitzkin (1989) and the power of British social constructionism (Bury 1986, 1987; Nicholson and McLaughlin 1987), but most American research on these social psychological topics has remained "safe" and ingrown as it has in most countries (Cockerham 1983, p. 1514; Claus 1983). In Britain and Germany, medical sociology has succeeded in becoming a required course for physicians but then ends up teaching them little more than social epidemiology and the social psychology of medical practice. Even the larger world of health activities and illness behavior not tied to medicine is ignored (Idler 1979). Although medical sociology was Gouldner's (1970) prototype of government-funded "theoryless empiricism," this largest of subfields is probably not alone in having a specialty journal where "there was no mention of any sociological theory whatsoever" in 171 articles during a five-year period (Cockerham 1983, p. 1519).

Governments, foundations, and the medical profession like their sociology focused on pragmatic rather than theoretical questions, and when studies are done from an independent sociological perspective, the status quo institutions take offense (Mechanic 1990, pp. 88–89). Yet I would contend that both grantors and the profession are best served by the very power and freshness of an independent sociological perspective, because it makes observations and produces insights—however uncomfortable—that are extremely helpful. This opinion revises the view that basic research in sociology is not compatible with applied or policy research

because it starts with different questions and often involves immutable variables such as age or class. That view is true as far as it goes, but I would contend that a basically *sociological,* "useless" analysis is worth its weight in policy gold, because it explicates what practitioners often ignore or take for granted.

No better example exists than what appears to be the worst example— the sociology of illness, illness behavior, the sick role, and stress. For even in its usual, apolitical form, the sociological view of these phenomena constitutes an alternate paradigm to the biomedical model embodied in most medical practice, a paradigm that challenges because it is more powerful, timely, and humane. Biomedicine has a few more cards to play, such as implants, transplants, pumps, and genetic engineering; but already it faces a clientele coping with increased dysfunction or pain. Their age, sex, class, ethnicity, religiousness, social networks, and community membership powerfully affect how they cope and how good medicine should be practiced. For, upon creative reflection, these social aspects of patients are not "immutable." Thus, I would urge that those who work with the sociology of illness boldly promote the power and relevance of the model of health care implicit in their research, particularly now when practitioners are so ready to receive it. Mechanic has put the case even more strongly, in terms of responsibility. An epidemiological, descriptive sociology "would simply be a servant to medicine, not fulfilling its larger responsibility to understand medicine as a social, political, and legal endeavor; to challenge its curative and technological imperatives; to examine equity of care in relation to class, race, gender, age, character of illness, and geographical area; and to study the appropriate goals and objectives for health care in the context of an aging society with an illness trajectory dominated by chronic disease" (Mechanic 1990, p. 89).

That argument and this issue, then, imply that developing medical sociology as an applied field runs the danger of making its work less useful, less powerful, and more vulnerable to takeover by adjacent disciplines than developing its distinctively sociological character.

The sociology *of* medicine, that is, the independent study of medical services, training, and institutions, has been more sociological in character. As Mechanic (1990, p. 87) has recently written, sociologists *in* medicine "work as applied investigators or technicians, seeking to answer questions of interest to the sponsors," while the sociology *of* medicine "focuses on testing sociological hypotheses, using medicine as an arena for studying basic issues in social stratification, power and influence, social organization, socialization, and the broad context of social values." A recent example is Samuel Bloom's (1988) institutional analysis of "reform without change" in medical schools over the past 30 years. Initial

attempts by defensive medical educators to suppress Bloom's paper were quickly swept aside by widespread appreciation for its insights, and it has shaped a major reform movement aimed at producing sociologically informed physicians. Bloom's study is an exception, however. Much of the sociology of medicine is vulnerable to being subsumed by health service researchers and health economists trained in schools of public health or management. Again the more service-oriented and less distinctively sociological this research, the more likely it will be subsumed.

Such exceptions aside, the question arises, How distinctive is the work done by most sociologists of medicine, and to what extent are they doing themselves in by not drawing upon or contributing to the larger field on a regular basis? For medical sociology, like the sociology of religion, of education, of crime, of the family, of work, of law, and of any social arena, has little theory of its own. Fortunately, this side of medical sociology can draw on the sociology of organizations, of the professions, of political theory, and of the welfare state. But how much do medical sociologists draw on this work, and how much momentum does this work have? Many of the organizational studies in medical sociology are technically good, but too often they do not bother to grapple with larger theoretical issues that would connect them to the larger discipline and also make them more consequential for practice. The number of medical sociologists who do research on health care systems in the sociohistorical tradition of Henry Sigerist (Falk 1958; Rosen 1958), or Bernhard Stern (Bloom 1990),[2] or Odin Anderson is quite small.

THE CONTRIBUTIONS OF THIS ISSUE[3]

The articles here each do one or more of five things: (1) use social behavior in health care to think through theoretical issues of general interest to the discipline; (2) use classic sociological variables and literature to gain new insights or make important observations about social behavior; (3) ask new questions and open up new arenas of sociological inquiry; (4) revise important concepts or observations long accepted in the field; and (5) document fundamental changes in health care as a social institution. The articles by no means represent all work in the field, or even all new and important work. For example, none represents research on women's issues, or the organization of technology, or the cultural organization

---

[2] Robert Straus said, "I consider Stern the unchallenged 'father' of medical sociology" (as quoted in Bloom 1990, p. 23)

[3] The articles are ordered by my judgment of what would create an interesting sequence of topics.

of wellness, or the management of chronic disorders. The list could be extended, but it does not detract from the offerings before us.

Sociological Implications of Regulation and Markets

The context for much new work is the relentless increase in medical services since Medicare and Medicaid were passed in 1965, and the responses to that increase. What Starr (1982, p. 420–49) characterized as "the coming of the corporation" actually consisted of two distinct, antagonistic developments: the rapid expansion of health care corporations as a natural if unwanted consequence of the protected markets that providers had established for themselves, and the intrusions by corporate or governmental purchasers to restrain health care providers (Light 1991). Mary Ruggie's article analyzes the effects of the largest institutional buyer: the U.S. Congress and its administrative agent, the Health Care Finance Authority. All they wanted was to cap Medicare expenditures for hospital care; yet soon they found themselves entangled in a labyrinth of state interventions. Moreover, their initially simple actions set off a chain of responses that restructured large parts of the health care system. Ruggie analyzes the underlying dynamic of what happened, and in the process she raises questions about the changing nature of the state.

Lee Clarke and Carroll Estes use new panel data gathered by Estes's research team to test the explanatory power of sociological and economic theories of markets. Like Ruggie, they note the "cascade effects of prospective payment," and their focus is on home health care agencies, an institutional population that has arisen in the post-Medicaid era. Because they are smaller and more fluid than hospitals, these agencies and similar new institutional populations offer sociologists new opportunities to use health care to test theories of organizations, markets, and networks. Of particular interest here is the economic and legal theory about how nonprofit organizations will behave, and the authors find little evidence to support it. This article, then, documents fundamental changes in society, uses data to test the validity of general theories, revises important concepts in the field, and opens up new arenas for future research.

Allen Imershein, Philip Rond, and Mary Mathis examine what happened to fragment the impregnable power elite that Alford (1975) found could block any health care reform. In doing so, they focus on the *disruptions* of network relations and the entry of non-health-care elites who altered the alignments both within the sectors of the health care elite and between them. The authors identify four kinds of market structures that result from these cleavages and examine some of their effects. Their analysis, like that of Clarke and Estes, maps new territory by focusing

on states rather than on federal policy; like Ruggie, they document the failure of incrementalism. One is tempted to add that the failure of incrementalism in American politics leads to still more incrementalism.

Professional Power and Professional Work

The sociology of the professions has long been influenced disproportionately by studies of the medical profession. This is particularly evident in the influence that Eliot Freidson's (1970a, 1970b) ideas about professional dominance have had over the past 20 years. Yet those ideas and observations reflect a certain historical period, which is passing, and the central emphasis on autonomy is becoming less tenable (Light 1988a, 1988b). Dissatisfactions have grown, but none of the alternatives has proven any better (reviewed in Light and Levine [1988]). Only Abbott (1988) has worked through a more comprehensive, historically informed model in *The System of Professions*. Sydney Halpern draws upon her comparative historical research to revise and extend Abbott's theory. She examines the changing relations between four medical specialities and their subordinate/competing occupations to show what "professional dominance" means and what factors affected its relative success.

The article by Martha Gerrity, Jo Anne Earp, Robert DeVellis, and me examines the concept of uncertainty, a quite different aspect of professional work long of interest to sociologists. We begin with a paradox of advanced industrial society: uncertainty expands with scientific and technical advances. Our review of socioanthropological studies indicates the potential of this area for future research, and the quantitative analysis of uncertainty shows how this elusive social phenomenon can be measured. Our findings about perceived uncertainty among physicians in practice challenges the main conclusion of sociological work over the past 15 years.

Theoretical Implications of Health Behavior

The final set of articles draws upon health behavior to make original contributions to sociology as a whole. Ellen Idler and Stanislav Kasl have written an article that looks at the religious factor. Beginning with an important review of Durkheim and his maturing view of religion in social life, the authors set out to show that religiousness is more than a matter of healthy life-styles, social contacts, and networks. They use panel data to document how religious involvement sometimes plays a protective role and other times plays a therapeutic role over and above its contribution through social correlates. Their study thus reaffirms the power of symbol and ritual in the face of uncertainty, pain, and fear.

Bernice Pescosolido has written the purest example of using health data to advance general sociological theory. Her article attempts to develop a more embracing theory of action than rational choice theory, using data about how people seek help when they are ill. Sociologists outside the specialty should note that the United States is gathering very interesting sets of data about attitudes, behaviors, networks, and organizations centered around health care that can be used to examine larger theoretical issues. Pescosolido's theory of action illuminates not only what actions people take but also how they do so through, in her words, "socially patterned networks of interaction producing systematic structures that link micro and macro levels" (p. 1126).

Thomas LaVeist's article stems from his prizewinning dissertation and opens a new area of research into the connections between political power and health status. While this idea can be traced back to Frederick Engels's treatise on oppression and illness ([1845] 1973), LaVeist approaches it with dispassionate, quantitative rigor that invites researchers in political and urban sociology and in public health to make it a regular part of their studies. Drawing upon a large population of cities and controlling for a number of possibly intervening variables, LaVeist finds an inverse relationship between a city's level of black power and the level of the black postneonatal mortality rate. He thus "extends present theory by demonstrating that not only do social factors such as poverty affect mortality, but the political organization of society also has important consequences in determining who lives, who dies, and when" (p. 1092).

SOCIOLOGY: GENERATIVE YET VULNERABLE

Sociology is the only discipline that systematically examines the structural and institutional forces that shape everyday life, behavior, and values. Other disciplines may do so to a limited degree, but only sociology makes such an analysis in all its manifestations the central focus of its work. Sociology quite often assumes a critical stance that allows greater dispassion and intellectual power than is often found in related disciplines. One could make the case that in its scope it leads the social sciences.

Yet sociology, in the United States at least, suffers from the constant raiding of its "stuff," and one reason may be the self-isolation of specialties. Over time, this makes them less sociologically distinct, more "neuter." We need to examine the degree to which this is a factor in the loss of political sociology or social psychology, which our discipline created and once dominated. The high regard for the *Annual Review of Sociological Methods* by other disciplines may be part of this same phenomenon. Part of the problem too is that imitation and adoption of our generative discipline is not seen by deans and other influential figures of higher

education for what it is: a high form of flattery and evidence of its contributions to many of the social sciences. Sociology therefore needs to trumpet what now feels like "raiding" as generativity through a series of well-placed articles, each coauthored by a prominent figure in a discipline that has drawn sustenance from sociology over recent decades. In the meantime, efforts like this collection of articles help in another way, by strengthening the intellectual ties between a specialty and the discipline as a whole.

REFERENCES

Abbott, A 1988 *The System of Professions*. Chicago: University of Chicago Press.

Alford, R R. 1975. *Health Care Politics*. Chicago: University of Chicago Press

Bloom, S W. 1986. "Institutional Trends in Medical Sociology." *Journal of Health and Social Behavior* 27:265–76.

———. 1988. "Structure and Ideology in Medical Education. An Analysis of Resistance to Change." *Journal of Health and Social Behavior* 29:294–306.

———. 1990 "The Intellectual in a Time of Crisis: The Case of Bernhard J. Stern, 1894–1956 " *Journal of the History of the Behavioral Sciences* 26:17–37.

Bloom, S. W., and R. E. Zambrana 1983. "Trends and Developments in the Sociology of Medicine." Pp. 73–122 in *Advances in Medical Social Science*, vol. 1. Edited by J. L. Ruffini New York: Gordon & Breach.

Bury, M R. 1986. "Social Constructionism and the Development of Medical Sociology " *Sociology of Health and Illness* 8:137–69.

———. 1987. "Social Constructionism and Medical Sociology. A Rejoinder to Nicholson and McLaughlin." *Sociology of Health and Illness* 9 439–41

Claus, L. M. 1983. "The Development of Medical Sociology in Europe." *Social Science and Medicine* 17:1591–97

Cockerham, W. C 1983. "The State of Medical Sociology in the United States, Great Britain, West Germany and Austria· Applied vs. Pure Theory " *Social Science and Medicine* 17:1513–27.

Collins, R. 1986. "Is 1980s Sociology in the Doldrums?" *American Journal of Sociology* 91 1336–55

Dahrenhorf, R. 1958. "Out of Utopia. Toward a Reorientation of Sociological Analysis." *American Journal of Sociology* 65:115–27.

Engels, F. (1845) 1973. *The Condition of the Working Class in England in 1844*. Moscow: Progress

Falk, L. A. 1958. "Medical Sociology. The Contributions of Dr. Henry E. Sigerist." *Journal of the History of Medicine* 13:214–28.

Freeman, H. E., and L. G. Reeder. 1957 "Medical Sociology: A Review of the Literature." *American Sociological Review* 22:73–81

Freidson, E. 1970a *The Profession of Medicine*. New York: Dodd, Mead.

———. 1970b. *Professional Dominance*. New York· Atherton.

Gerhardt, U. 1989. *Ideas About Illness· An Intellectual and Political History of Medical Sociology*. London. Macmillan

Gouldner, A. 1970. *The Coming Crisis of Western Sociology*. New York: Basic.

Idler, E. 1979. "Definitions of Health and Illness and Medical Sociology." *Social Science and Medicine* 13A 723–31

Light, D. W. 1988a "Social Control and the American Health Care System." Pp. 456–74 in *The Handbook of Medical Sociology*, 4th ed. Edited by H. E. Freeman and S. Levine New York. Prentice-Hall.

————. 1988b. "Turf Battles and the Theory of Professional Dominance." *Research in the Sociology of Health Care* 7:203–25.

————. 1991. "The Restructuring of American Health Care." Pp. 53–65 in *Health Politics and Policy*, 2d ed. Edited by T. J. Litman and L. S. Robins. New York: Delmar.

Light, D. W., and S Levine. 1988. "The Changing Character of the Medical Profession· A Theoretical Overview." *Milbank Quarterly* 66, suppl. 2.10–32.

Mechanic, D. 1990. "The Role of Sociology in Health Affairs." *Health Affairs* 9, no. 1:85–97.

Mills, C. W. 1943. "The Professional Ideology of Social Pathologists." *American Journal of Sociology* 49.165–80.

Nicholson, M., and C. McLaughlin 1987. "Social Constructionism and Medical Sociology: A Reply to M. R Bury." *Sociology of Health and Illness* 9:107–26.

Rosen, G. 1958. "Toward a Historical Sociology of Medicine: The Endeavors of Henry E. Sigerist " *Bulletin of the History of Medicine* 32:500–516.

Starr, P. 1982 *The Social Transformation of American Medicine*. New York. Basic.

Straus, R. 1957. "The Nature and Status of Medical Sociology." *American Sociological Review* 22:200–204.

Waitzkin, H. 1989. "A Critical Theory of Medical Discourse. Ideology, Social Control, and the Processing of Social Context in Medical Encounters." *Journal of Health and Social Behavior* 30:220–39.

The Paradox of Liberal Intervention: Health Policy and the American Welfare State[1]

Mary Ruggie
Columbia University

This article presents the Medicare and Medicaid programs as typi-
cal examples of liberal interventionism and then discusses the con-
sequences of present patterns of state intervention in U.S. health
care. Before 1980, the U.S. government assisted both providers and
recipients through financial outlays. Now, however, while seeking
to cut costs, the government has become more involved in regulat-
ing the activities of both hospitals and physicians. The vehicle of
this change in state intervention is the diagnostic related group, or
DRG, system of prospective payment to hospitals, adopted in 1983.
The article shows how the DRGs affect the delivery and substance
of health care and concludes that DRGs enhance the capacity of
the state to govern the health care system as a whole and manage its
factors of production. This "step-by-step" growth of government
regulation in a political atmosphere that endorses private enterprise
and marketplace rate setting is termed the "paradox of liberal inter-
vention."

There is no doubt that over the last decade or so health policy in the
United States has moved from an area beset by "dynamics without
change" (Alford 1972) to one "exhibiting both colorful dynamics and
genuine change" (Brown 1986, p. 571). New departures are appearing
in every aspect of health policy, including the delivery and organization
of care, the status of the professions, and the role of both the public and
private sectors. Health care is unfolding as a unique locus of reform

efforts in social policy, outpacing developments in other fields, and as a result, it offers a glimpse into the future of the American welfare state.

While it is clear that a transformation of some sort is underway in health policy, no consensus has yet emerged on the full significance and meaning of the changes, let alone their concrete implications for welfare state development. Some observers lament the continuing drain on the economy posed by the huge expenditures of an unwieldy medical system (Thurow 1985). Others decry the "significant reductions in eligibility and benefits" resulting from decreasing government contributions to total health care expenditures (Mechanic 1986, p. 34). Some suggest that the appearance of large corporations in American medicine signals "the failure to rationalize medical services under public control" (Starr 1982, p. 449). Others observe that government-subsidized health insurance for the middle classes has created a semiprivate welfare state (Rothschild 1988).

These diverse assessments reveal both the complexity of the issues as well as the difficulty of finding a coherent framework within which to analyze them. Moreover, they all imply if not a retreat by government from its responsibility for health care, then at least a failure by government to construct an adequate system of health care. Permeating these analyses of change is a pessimism about the future of American health care in an era of government retrenchment and cost containment. Such pessimism feeds into growing concerns about the future of the American welfare state. How can we expect to continue to provide adequate services, we ask, in light of recent cutbacks in social provision? (See Brown 1988; Weir, Orloff, and Skocpol 1988.) My response is that the role of the state is changing in important ways that, far from diminishing the American welfare state, will deepen and eventually expand it.

This paper thus offers a very different interpretation of recent developments in health policy. It suggests that, by viewing the health care sector through the broader lens of changing forms of state intervention in the provision of human services, developments in health care come to reflect what I shall call "the paradox of liberal intervention." Specifically, though present changes in state intervention were designed for the purpose of reducing the fiscal exposure of the state and limiting state involvement in the delivery of care, state intervention in health care in fact is evolving toward a more intrusive system of management. The previous system of health care, in which government simply assisted both the providers and the recipients through financial outlays, is giving way to a system in which government is prepared to regulate more fully the activities of both the owners of health care facilities and the producers of health care services. The issue is not only what government is *doing*, however, but also what developments government is *inducing* among the providers of health care. Paradoxically, changes in the form of state

intervention are altering the structural bases of health care toward a system that may yet lay the foundations for a future national health insurance system, albeit one suited to the unique features of the American sociopolitical framework (Morone and Dunham 1985). Through its exercise of fiscal control, the state is acquiring a capacity to govern the health care system as a whole and manage its factors of production. The earlier absence of this capacity characterized American exceptionalism in the provision of health care. Its emergence suggests the end of American exceptionalism in this area and underpins future evolution of the American welfare state.

To elaborate this argument I examine the intended and unintended health care policy consequences of the 1983 legislation that authorized a new method of reimbursement in Medicare and (where applicable) Medicaid. This decision changed financing from retrospective cost-based reimbursement to prospective payment based on diagnostic related groups or DRGs (Fetter et al. 1980; Morone and Dunham 1985). The DRGs set a fixed price for treating categories of illnesses; the objective of the system is to control costs and thus reduce government funding of health care. In several of its features, the DRG system contradicts the format that commonly characterizes new developments in American social policy (Weir et al. 1988). It did not emerge through political class struggle or group coalitions, or from a Democratic administration, or local or state regulatory laws. Instead, it arose from a proclaimed effort on the part of the federal government to minimize its regulatory intervention by acting as an institutional buyer (Light 1991). In contrast to the open-ended system of reimbursement of costs in the original Medicare and Medicaid programs, the closed system of prepayments under the DRGs constrain provider choices about the level and kind of health care. Ironically, through the minimalist interventionism of the DRG system, the state is performing for the first time two critical activities: controlling the main social actors in health care provision (hospitals and, indirectly, physicians) and coordinating this area of social expenditures with broader economic objectives. The result, I contend, is a strengthening in the fundamental foundation of the welfare state.

LIBERAL INTERVENTION AND HEALTH CARE

Ambiguity about the status of the United States as a welfare state runs deep throughout the relevant literature: the United States is typically characterized as the "richest" of the advanced industrial societies and yet the "most laggard" in developing social welfare measures (Wilensky 1975). Many studies explain American exceptionalism in terms of both the antecedents to and the outcomes of welfare state development (Kudrle

and Marmor 1981; Quadagno 1987; Amenta and Skocpol 1989; Skocpol 1988). As long as the emphasis remains on the level of state expenditure, and the quantity and scope of government programs for social services, then the United States does remain an outlier: the percentage of GNP devoted to welfare expenditures tends to be lower, and its programs less universal, than those of several other leading welfare states—though not all, it should be noted.[2] And as long as theorizing about the welfare state remains dominated by concerns about the size of the public sector alone, then the United States does posit an aberrant case in welfare state development (Starr 1989). But these views are incomplete, for the relationship between the public and private sectors must also be considered.

The United States has been a "liberal" welfare state, less inclined to adopt centrally directed intervention than its European counterparts.[3] A liberal welfare state seeks to intervene minimally in a market economy; intervention is undertaken only in order to compensate for adverse consequences of private or market forces. State intervention in a liberal order is intended neither to supersede market forces nor to dictate or direct the conduct of private affairs, merely to enable them to return to a condition of "normalcy" (Gutmann 1988; Moon 1988). Thus, keeping the domains of state and society distinct is fundamental to a liberal order.

However, certain anomalies may arise in the attempt to maintain this separation of state and society. There are two possible sources of anomalies. One is the disposition of a liberal welfare state to limit its sphere of intervention in favor of the market, which requires strict specification of the proper boundaries of state intervention. The accumulation of these specifications itself frequently results in deep forms of intervention. Another source of anomaly is the private sector. In its capacity as a provider of services, the private sector may try to outwit or manipulate the government as a monopsony buyer of those services. When such behavior occurs, the state in turn is drawn into deeper forms of intervention as it attempts to limit its vulnerability to such exploitation and to correct the adverse social consequences of such private sector behavior. The actual extent of intervention and the forms it takes will differ, depending on

[2] "For instance, in 1985 total expenditure for health care in the United States, measured as a percentage of GDP (10.8%), exceeded that in Britain (5 7%) and Canada (8.6%). Total health spending per capita also placed the United States ($2,354) higher than Britain ($836) and Canada ($1,683). However, the public-sector portion of these expenditures in the United States (41.1%) was much lower than in Britain (90.0%) and Canada (75.9%) (Schieber and Poullier 1991).

[3] See Esping-Andersen (1989) for a discussion and typology of welfare states and policy regimes

the sources of the anomalies to which the state is responding.[4] These variations can occur across sectors within a single welfare state or across welfare states within a similar sector of social provision.[5] The developmental path of a liberal welfare state thus constitutes a paradox: more than intended intervention frequently is the product of the very attempt to limit intervention.

In the case of the DRGs, this paradox is at the point of producing transformational consequences. The DRG system is allowing the state to acquire a critical capacity that it has always lacked, namely, the ability to control those social actors whose demands and interests it must also mediate. To demonstrate this argument, it is necessary first to reexamine the nature of state intervention in the original Medicare and Medicaid programs.

The role of the state in health policy prior to the 1980s was symptomatic of the American approach to the development of the welfare state.

---

[4] Examples of deeper than intended intervention in the United States abound in the fields of social welfare and beyond. For instance, in carrying out the stringent eligibility criteria of the Aid to Families with Dependent Children program, U S government social workers, in order to determine that there was "no man in the house," so intruded into the private lives of recipients that they undermined the very possibility of a surrogate for normal family life (Piven and Cloward 1971). The nursing home industry offers an example of the consequences of minimal regulation in the private sector that required ever greater intervention, including government funding, to correct for poor conditions and the mistreatment of patients in the provision of care (Vladeck 1980). Yet another example appears in recent legislation that seeks to tie the receipt of welfare to participation in the work force, an effort that will necessitate greater involvement by the state in the labor market than has existed up to now. Ironically, the measures are intended to reduce government's presence in welfare payments Also, Smith and Stone (1988) show that the practice of contracting out to private agencies for the performance of public welfare functions does not reduce the role of government Instead, it "creates a new set of interests for public spending," builds pressure for "greater central management and regulation," and leads "to a merging of the public and private sectors rather than a simple transfer of funds from one to the other" (p. 247). In a very different area, it is interesting to note a *Wall Street Journal* complaint that congressional involvement in defense contracts also exhibits the pattern of increasingly deeper intervention than intended: as a result, the cost of approved weapons systems appears to have been raised by as much as $30 billion per year (Melloan 1988).

[5] Canada and Britain are two liberal welfare states similar in many respects to the United States. Instances of the paradox of liberal intervention occur in both countries, especially in the field of social welfare (see Ruggie 1984, chap. 5). But in the field of health care, the responses of the private sector and the medical profession to state intervention are more amenable, and the insistence of the state on maintaining the state/society boundary is more lax than in the United States. Consequently, the outcome of state intervention in formulating a system of national health insurance in both countries is planned, not an unintended consequence

The principles of both the Medicare and Medicaid programs, as formulated in 1965, typify the approach: minimal interference in the essentially market-based determination of access to and delivery of care, coupled with ad hoc, piecemeal intervention aimed at keeping the realm of public responsibility distinct from the private sphere and thereby limiting the extent of state intervention. However, the accumulation of interventionist measures eventually created a greater degree of intervention than originally intended in both programs. While themselves exemplary of the paradox of liberal interventionism, the Medicare and Medicaid programs essentially laid the groundwork for the ultimate paradox that ensued with the DRGs.

Medicare

Medicare is unique among federally administered social programs in the United States in that it is organized for a stable population—persons aged 65 and over and the disabled—that can be readily identified without resorting to intrusive means tests and that cannot lose its eligibility.[6] However, the role of the federal government in Medicare has been minimal. The very initiation of Medicare was considered so threatening that hospitals and the medical profession made sure its structure would not lead to a "socialized" form of medicine (Starr 1982, pp. 363–78; Light and Levine 1989). Rather than set a new course for health care financing, the government was constrained by the hospital industry and the physicians' lobby to follow the existing model of private insurance. The providers of health care, both hospitals and doctors, insisted on being reimbursed on a "reasonable" cost basis, including assistance for "all necessary and proper expenses" (U.S. Congress 1966). As a result, reimbursement was very generous, while standards and terms of reimbursement remained only loosely defined. As stated in the original legislation, which contained a "Prohibition against Any Federal Interference," the government minimized its role in regulating the delivery of health care (U.S. Congress 1966, sec. 1801). Those regulations that did exist concerned macro-level guidelines on what types of medical procedures and treatments were and were not reimbursable (mostly for physicians' services, not hospital costs) and what type of review procedures hospitals were to establish to monitor their services and expenditures. In other words, the state delegated to the providers of care much of the power that accompanies the purse. The overall structure of the initial Medicare

---

[6] The Veterans Administration program has comparable features. In contrast, eligibility for the Medicaid program is less consistent and more variable

program resulted in a passive form of governmental intervention leaving much discretion for program implementation to the providers of care.

It became obvious early on, however, at least by 1969, that the open-ended financing of Medicare was bringing about an inflation of health care expenditures. Therefore, in a second stage of intervention, aimed at bringing costs under control, the government began to develop such measures as hospital rate-setting, certificate of need requirements, and professional standards review organizations. Each of these programs themselves constitute an example of how limited measures give rise to ever greater degrees of intervention. For instance, certificate of need regulations and health systems agencies were attempts to rationalize planning for health care facilities and expenditures, but were sufficiently fragmented to require incremental adjustments that merely complicated the efforts without creating any restraints on rising costs (Starr 1982, pp. 398–419). These measures, then, increased the regulatory mechanisms available to the government, but they did not rein in rising expenditures. For, while more and more burdensome administratively, the measures refrained in typical liberal fashion from interfering with the internal organization of provider-determined, cost-based care. By the end of the 1970s the accumulation of government regulations was declared a failure by all sides—the measures had accomplished neither cost containment nor reform in the medical system (Starr 1982, pp. 405–17). The conditions were thereby in place for the next stage of intervention, which began in the early 1980s.

Medicaid

A similar cycle of increasing intervention occurred in the Medicaid program. Medicaid is intended only for a selected group of individuals: the poor, specifically Aid to Families with Dependent Children (AFDC) recipients, and the institutionalized aged, blind, and disabled. Intergovernmental relations have been organized in such a way that the federal government contributes about one-half to three-quarters of the states' costs for the program. The amount has varied over time, depending on the state and the fiscal goals of the federal administration. The federal government has imposed two main requirements for the receipt of these funds. First, states must provide coverage for five basic health services to certain categories of persons. Second, states must reimburse hospitals in terms of the same open-ended, cost-based, provider-determined reimbursement as in the Medicare program. When it came to regulating the *providers* of health care in the early stages of Medicaid, the role of government at both the federal and state levels exhibited some of the same features of hesitant and incremental intervention characteristic of the

early Medicare program. For example, federal authorities gave few guidelines on how states should provide the required services and transferred responsibility for most of the administrative tasks to the states. However, the requirements for hospital reimbursement led to rising expenditures and severely constrained the capacity of states to provide the required services. The federal structure of government relations therefore indirectly maximized the power of providers, in that it left individual states with insufficient political and economic power to develop effective regulatory measures (Vladeck 1979).

In order to correct the problems they were experiencing in their relations with providers, state governments began to adopt regulatory measures, embarking on a cycle of increasingly deeper intervention similar to that seen in the Medicare program. For instance, regional and state-level health planning agencies blossomed throughout the 1970s. These agencies strengthened state leverage over providers by elaborating the bureaucratic edifice of health care. But they did not improve efficiency in resource utilization nor achieve cost savings. Hence, in a further effort to control resources, Congress passed the 1931 Omnibus Budget and Reconciliation Act (OBRA), which gave states additional power over providers by allowing states to limit choice of provider for Medicaid recipients and to require that Medicaid recipients use health maintenance organizations (HMOs) if the states so chose. The inadequacy of earlier measures thus led to stronger efforts in regulation.

When it came to the *recipients* of care, Medicaid evidenced even greater administrative activism by both the state and federal levels. Measures aimed at maintaining the limited sphere of state intervention have instead resulted in increasing levels of intervention. In 1965 each state devised its own means tests to determine eligibility for Medicaid. Incomplete and variable as these tests were, additional legislation necessarily accumulated to take into account changes and ambiguities in the status and therefore eligibility of recipients.[7] The determination and implementation of these criteria have been arbitrary, with vast differences existing among states in their discretionary definition of "medical indigency" (Cromwell et al. 1987). Cutbacks designed to reduce government expenditures commonly focused on withdrawal of optional services or greater restrictions on access to optional services. Since these constituted only

---

[7] For instance, methods were developed to meet, temporarily, the continuing health needs of families who have managed to cross the income boundary but still cannot buy insurance (generally some additional months of provision is allowed), as well as the constant needs of children—including the unborn children of pregnant women—in poor families who are not eligible for AFDC (usually because they are members of two-parent, low-income families)

minor reductions in state budgets, the cutbacks burdened the recipients of Medicaid without achieving significant cost savings. Additional measures, devised either to compensate for inadequacies in the program or to reinforce previous regulations, grew in number and eventually resulted in a vast array of interventionist instruments.

The accumulation of government measures created a cumbersome system of medical provision, which nevertheless still failed to reach one-quarter of the poor. Moreover, it failed to meet adequately the health needs of those it did reach, especially in terms of preventive care, thus requiring still additional methods of health care provision. And yet, for all the intrusiveness into the lives of the recipients of Medicaid, the chief regulatory features of the program are concerned primarily with the extent of coverage and assistance; they make little reference to the health of the recipients of care or to the quality and character of the delivery of care. While physicians' practices are not the object of intervention, treatment of Medicaid patients is nevertheless indirectly affected by statutory limitations on Medicaid reimbursement, which in some states fall below 50% of physicians' fees (Perloff, Ketke, and Neckerman 1987). To put it in analytical terms, the main purpose of the intrusiveness in Medicaid is to safeguard the strict boundary between the eligible and the noneligible, and between public and private, an emphasis on separation that is typical of liberal interventionism. Beyond that, the orientation of Medicaid, as well as Medicare, has been to defer to the providers of health care in their decisions on the kind and extent of provision and to reimburse them up to the statutory limits.

In their very attempts to limit state intervention, both the Medicare and Medicaid programs have ironically resulted in deeper intervention than intended. They have both expended a great deal of effort on maintaining the boundaries between state and society, public and private. Their forms of intervention and their lack of coordinated regulation complicated the system of health care without violating its structural premises. Yet these interventions also drew government more deeply into dealing with the problems of cost and therefore of provision and delivery of services. In the process, the unmanaged approach of liberal interventionism has inadvertently laid the foundations for a more managed approach, resulting in a more profound paradox in the development of the American welfare state.

DRGS AND THE DEVELOPMENT OF HEALTH CARE

Throughout the 1970s, under both Republican and Democratic administrations, the government embarked on a cycle of increasing intervention in several facets of health care in an attempt to control costs. By the

late 1970s, however, it was widely recognized that these efforts had not contained health care expenditures. And so, by the 1980s, with a new administration in power professing, indeed pressing, a new philosophy of less government in favor of more private sector competition, a new stage in health care policy commenced. In 1982, convinced of the need to directly cap government reimbursement, Congress passed the land- mark Tax Equity and Fiscal Responsibility Act (TEFRA), which paved the way for prospective payment. Congress further instructed the Depart- ment of Health and Human Services (DHHS) to review and present for consideration various cost containment demonstration studies and research projects that its branch agencies had been funding and oversee- ing for several years. Finally, in 1983 Congress adopted the model recom- mended by the Health Care Financing Administration (HCFA) for imple- mentation in Part A of the Medicare program—that is, for payment to hospitals.

These congressional actions in the early 1980s were a mark of frustra- tion with the government's past inability to contain rising costs in health care. They therefore constituted new attempts to develop measures that would explicitly limit the fiscal and administrative roles of government by shifting more of the burden onto the private sector. The DRG system was considered to be a conservative measure to contain expenditures, one which would simply cap the government's contribution and free the private market from further regulation, thereby promoting competition among providers. Moreover, Congress legislated the reform without ma- jor opposition or input from the primary interest groups.[8] The American Medical Association (AMA) was surprisingly diffident about the measure, probably because it was not intended to affect physicians. And hospital lobbies in general supported the new legislation "because it appeared to pose fewer threats to their revenues than other measures under consider- ation" (Thompson 1987, p. 720). The ease with which the legislation passed and the absence of fanfare about it is ironic given its subsequent effects.

The new system of prospective payment replaced the previous system of retrospective reimbursement in which providers were reimbursed for costs after services had been delivered. Prospective payment is based on a preset schedule of payments, determined before services are delivered. Under the DRG system hospital payments are based on approximately 470 groupings of diseases or illnesses (Fetter et al. 1980). Ideally, physi-

[8] The mental health lobby is an exception Representatives from the major interest groups testified about the inappropriateness of the mental health DRGs and received an extension for the full implementation of the categories covering mental disorders. The substance-abuse DRGs have since been implemented (Ruggie 1990).

cians in hospitals diagnose a patient's ailment at the time of admission, consult the schedule for the level of payment assigned to the ailment, and then treat the patient knowing what the maximum cost can be. In many cases hospital administrators assign the DRG category according to the physician's diagnosis and then monitor the cost of treatment. An incentive to control costs is built into the system because if hospitals can treat the patient for less than the payment level, they can keep the difference. If treatment exceeds the payment level, the hospital is responsible for the difference. In adopting the DRG system the government shifted its health care funding approach from that of a passive payer to one of an active buyer. The DRGs constitute, then, a mechanism devised by a cost-conscious buyer to force uncontrollable complex purchases of inpatient hospital care for the elderly into fixed price transactions.

In its first phase of implementation the DRG system was to be "budget neutral" (Director, PPAC, personal interview, August 1988). It was touted as a market mechanism enabling hospitals to correlate cost and price. The DRG system uses a statistical averaging of past medical experiences and expenditures under retrospective reimbursement insurance. Each standardized diagnostic coding is given a weighting which reflects a match between past average costs for treating the disorder and the government's payment level. Several specific measures of flexibility are built into the system to reduce its arbitrariness given the diversity of hospitals and treatment situations across the nation.[9] In addition, both the rates and the categories are subject to recalibration as new information materializes on the DRGs' ability to capture the match with costs. The DRG system began with a simple intention. But it has become the foundation of a radical restructuring of health care financing.

The consequences of implementing the DRG system are proving to be nothing short of profound. Having to work within averages represents a critical constraint for health care providers. Hospitals are expected to conform to the standardized categories when recording diagnoses and to provide treatment within the aggregate averaged costs; physicians are expected to cooperate with their hospital's administrative needs and to make medical decisions about inpatient care for the elderly that maintain the hospital's solvency. Moreover, there is no doubt that the overriding purpose of the DRG system is cost containment. The government intends for hospitals to operate more efficiently and for government expenditures

---

[9] For instance, adjustments exist for different levels of expenses—wage costs in particular—for hospitals in urban and rural areas, as well as for teaching and non-teaching hospitals And a so-called disproportionate share adjustment favors inner city hospitals with a large number of indigent cases

K-P-12

on health care to be capped, a combination that is designed to effect a slower rate of growth in overall national health care expenditures. As providers learn how to operate within the set limits, further reductions become possible. Therefore, the immediate constraint of averaging also contains long-run expectations for behavioral changes (DHHS/HCFA 1985). All of these constraints imply a shift from the original basis of the DRG system—that is, a consensus-based, adjustable, and flexible determination of inputs and prices—to increasingly more unilateral management of prices on the part of the state. And in order to achieve its straightforward goals, the government has had to expand its bureaucracy and its system of regulations to monitor health care provision and respond more effectively to the infinite complexity of pricing. In short, the consequences of implementing the DRG system signal the onset of another stage in state intervention in health care, belying its initial minimalist rationale.

There still remain considerable disputes about the appropriateness of the DRGs (see, e.g., a special issue of *The Journal of Medicine and Philosophy,* edited by Stuart F. Spicker, May 1987). However, even a cursory examination of the relevant literature reveals that very few practitioners or analysts dismiss the DRGs as inconsequential. It is not my intention to review the technical or ideological debates about the DRGs. Rather, my purpose is to use these debates as data points to elaborate the potential consequences of the DRGs. In and of themselves the DRGs may be a simple pricing system, but the imposition of DRGs exerts state influence over health care and its factors of production as never before. Health care providers, in turn, are adapting to the state's agenda of cost containment. The following discussion demonstrates that the implementation of the DRGs is fundamentally affecting the substance and delivery of health care, symbolizing another layer in the paradox of liberal interventionism.

Categorization and Standardization

Diagnosing a patient's ailment in terms of commonly used categories is everyday practice for physicians. The task of code assignment ought, then, to pose no inherent problem. But under the DRG system codification is not merely a medical exercise; it is also a financial task with management implications. As a result, codification has raised problems for both physicians and hospital administrators. Since each category carries a relative weight for reimbursement, it is important that the choice of diagnostic category conform to the cost of possible treatments. Take, for example, the following clinical presentation: "patient is brought into emergency ward unresponsive and with depressed respirations; gastric

aspirate shows traces of barbiturates" (Iezzoni and Moskowitz 1986, p. 928). The attending physician or hospital administrator can choose to categorize the patient in DRG 449 (toxic effects of drugs at age over 69 years and/or complication or comorbidity), which carries a relative weight of 0.7331, or the choice can be DRG 23 (nontraumatic stupor and coma), which carries a relative weight of 1.1568. The former category describes the more precise clinical diagnosis, but carries a 40% lower level of payment than the latter category, which describes the symptom less rigorously.[10] Once the physician or administrator knows the payment implications, the choice seems to be clear.

And in case it is not, other administrators or hospital review boards monitor categorization decisions to ensure that revenue-maximizing choices are made. The government is beginning to recognize the full implications of "DRG creep" and is responding to the profits that hospitals have gained at its expense by stepping up its mechanisms of control. The HCFA, which is responsible for monitoring and correcting the problem, has engaged in several efforts to readjust the weightings of categories.[11] Each reimbursement submission is reviewed by peer review organizations (PROs) which are under the central control of HCFA. They have added increasing sanction power to their vigilance in monitoring hospitals to check for unwarranted DRG creep and enforcing compliance with government requirements. These activities have engaged the government in far more intrusive interventions than simply setting prices as originally intended. And through this new kind of regulatory role, the government is becoming capable of managing the costs and outcomes of health care provision.

When first proposed, the DRG system consisted of 383 categories of illness or disease of the elderly, grouped by anatomic areas and organ systems. By the time the system was implemented in 1984, the number of categories had expanded to 467. Over the years a few more categories have been added to reflect diagnostic precision and to incorporate some illnesses of the nonelderly under the Medicaid program. The standardization in codification itself constitutes a remarkable achievement of the

[10] To cost out the difference, suppose the patient is hospitalized for three days at an actual cost of, say, $500 per day. Choosing DRG 449 yields a reimbursement of about $1,100; choosing DRG 23 yields about $1,735.

[11] For example, in the June 1987 update of the DRGs, the higher weighted category (DRG 23) was reduced to a value of 0.9438 and the lower weighted category (DRG 449) was raised to a value of 0.7904, leaving only a 15% difference between them. While in general more highly weighted categories are still being used more frequently, there is now more assurance from peer review organizations that the practice is the result of better coding procedures and better understanding of severity of illness (director, HCFA, personal interview, August 1988).

DRG system, given that before 1983 "there was no standard medical coding system for all Medicare carriers to use in paying for and reporting physician services" in hospitals (DHHS/HCFA 1985, chap. 6, p. 10). Still, physicians and hospital administrators regard the list as inadequate. The first few years of implementation produced considerable demand to improve the description of medical conditions in the categories and to accommodate expanding knowledge about medical disorders. One major problem is that no single DRG category contains an adequate mix of circumstances that can arise in a single episode of illness. The problem is not semantic, but pertains to the different costs of treating different episodes of illness that were initially placed within similar diagnostic categories. To appreciate the diagnostic constraint imposed by the DRGs we need only compare the number of categories with the thousands of diagnostic items in each edition of the *International Classification of Diseases*, the main reference manual for physicians. There can be little doubt that the intervention represented by the DRGs has been intrusive, constraining physicians and administrators in their everyday tasks. And because physicians must now take nonmedical factors into account, they are concerned that their medical prerogatives are interfered with by the new system.

   The categories were refined to take into account the criticisms of physicians and the experiences of clinicians. However, the task has not been left to clinicians alone. The government has a stake in keeping the number of categories as limited as possible. In its position as a monopsony buyer, the government is able to exercise a significant degree of power over the propensity of clinicians to expand the categories. Clinicians now regularly meet with government officials under the auspices of the HCFA to negotiate refinements in the categories. Changes are then sent to another branch in HCFA where recalibrations in the DRG weightings are calculated for payment purposes. Most of the change has occurred in the weightings not in the categories themselves. HCFA administrators are optimistic that a workable degree of homogeneity in resource consumption within the categories is being approached. There is now a general consensus that the DRG categories will continue to be refined to improve the case mix index and reflect clinical experience, but that further significant change in the categories will not be needed (director, HCFA, personal interview, August 1988).

   An important consequence of the implementation of and refinement in the DRGs is the unprecedented negotiations between clinicians and government officials on medical matters. It is ironic that clinicians are now collaborating on the details of state intervention into their own work. Physicians are learning to work within the DRG framework of standardized categories and to develop alternative methods of health care appro-

priate to the new limits being imposed on them. Hospitals are developing improved administrative techniques, including better coding procedures and record-keeping practices. Although the government's dominant interest in the negotiations over categories and weightings is financial, it has become, ipso facto, involved in the micromanagement of health care and the assimilation of health care delivery and provision to the new system of payment. The role of the state is thus betraying its own minimal interventionist, market-determined liberal model. And the outcome of the new mode of state intervention is a more managed system of health care, governed by more authoritative state-determined rules.

Hospitalization

The most clear and consistent impact of the DRG system thus far has been on the course of hospitalization. State intervention has led to a reduction in the length of hospital stay and has influenced decisions about admission to and discharge from the hospital. These are all critical judgments in an episode of health care, and they carry significant implications for further secondary consequences of the DRGs.

The first and immediate response of hospitals to the change from a per diem to a per discharge basis of payment was to attempt to reduce costs to within or below the DRG reimbursement level by shortening patient stays in the hospital. In the first year alone of DRG implementation the average length of stay of Medicare patients in short-term hospitals declined by 9%, from an average of 10 days for each episode of illness in 1983 to 9.1 days in 1984 (Guterman and Dobson 1986, pp. 103–4). Hospital stays were already becoming increasingly shorter throughout the 1970s; DRGs reinforced and accelerated the trend. For a few years the length of hospital stays continued to decline at a steady though less dramatic rate that was more reflective of the earlier trend (PPAC 1988, p. 11). The average length of stay has now leveled off. Physicians and hospital administrators complain about the pressure to discharge patients early under the DRG system. One AMA-sponsored survey found that "43 percent [of the physicians surveyed] reported that there was pressure to discharge patients early" (cited in Iglehart 1986, p. 1464). In contrast, before hospitals adopted the DRG system, physicians were more likely to accommodate patient preferences for longer stays and might even use longer stays to compete for patients (Robinson et al. 1988).

Earlier discharges in DRG hospitals have led to a wave of secondary consequences. In particular, the DRGs have resulted in more standardized patterns of discharge rates for major surgeries for all patients (not just the elderly) across hospitals. For example, the length of time spent in the hospital for acute appendicitis is now relatively similar across the

country despite differences in hospitals, physicians, and patients. And hospitals are routinely discharging patients earlier for minor surgical procedures, such as tonsillectomies.[12] DRGs have also influenced patient care after discharge. Some patients who are able to be sent home are receiving more continuous care on an outpatient basis, but a more prevalent occurrence among the elderly is sending patients who cannot return home to nursing homes.[13] It appears, then, that the DRGs are institutionalizing a trend in hospitalization that had been incipient for some time, situating hospitals as facilities for the short-term care of acutely ill patients.

One further change, in admission patterns, has occurred across as well as within hospitals. Government agencies were concerned that the DRGs would encourage increased hospital admissions since each extra admission generates additional revenue for hospitals. Admission of Medicare patients had been increasing before the implementation of the DRGs (by an average of 4.8% between 1978 and 1983). But the anticipated further increase after the introduction of the DRGs did not materialize; in fact, a decline occurred (by an average of 2.2% between 1984 and 1987 [PPAC 1988, p. 12]). The previous growth in admissions is one trend that the DRGs actually reversed. Although the decline in admissions is occurring under the guise of reducing costs in keeping with the objectives of state intervention, it is creating a situation that may contain unsettling consequences for the health care system and require yet further state intervention. Some of the decrease in hospital admissions is among patients who can be taken care of on an outpatient basis. Shifting the locus of their care amounts to a transfer of payment agent. However, some of the decrease in admissions is also among patients who are "financial losers" for hospitals, that is, patients who are seriously or chronically ill and whose treatment costs would therefore exceed the levels of payment under the DRG system. Physicians acknowledge that they are not admitting some patients they would previously have admitted because their care needs are too costly (Notman et al. 1987). These patients are, however, beginning to burden the health care system as a whole. Shifting them to public facilities, hospital or nonhospital, or to nursing homes relieves

[12] There are still large differences among hospital discharge rates for various medical disorders, however, reflecting continuing variations in physician practices (Lagoe 1986).

[13] One study comparing pre- and post-DRG discharges in a facility that used the psychiatric DRGs reported that, after DRG implementation, more psychiatric patients (both Medicare and Medicaid) were being sent to "more restrictive" facilities (other hospitals, nursing homes) rather than to their own homes, presumably because the patients were being discharged early and still required care (Thienhaus and Simon 1987).

some of the DRG hospitals but not the responsibility of the government and the Medicare and Medicaid programs. A paradox arises as the government finds itself reducing expenditures in the area of its intervention but facing the possibility of second-order rising expenditures overall, calling for yet further intervention.

The DRGs have clearly changed hospitalization patterns by reducing admissions and length of stay and creating greater overall uniformity. Thus, state intervention aimed at containing costs has become the vehicle for a significant change in one key dimension of health care delivery and provision. The changes in hospitalization are also significant for the ongoing history of medical care in the United States. For decades hospitals were the center of medical knowledge, practice, and treatment in the nation. Their eminent place and their pivotal role are now declining, accelerated if not caused by state intervention through the DRG system.[14] In place of hospital-led health care, the United States is developing a state-directed, managed system of delivery and provision—in Medicare at least for now.

Patterns and Location of Treatment

Even though the DRG system "was not designed to directly impact upon the physician" (DHHS/HCFA 1985, chap. 6, p. 8), hospitals are encouraging if not pressuring physicians to alter their treatment practices in accordance with the hospital's structured limitations. Enough physicians have complied that notable changes in treatment are emerging. Evidence is accumulating that the DRGs are affecting two aspects of treatment in particular: the intensity of services in hospitals and the location of care. It is becoming clear that the DRGs are functioning as a system of constraints that promotes tighter management of resources and, indirectly, reshapes certain components in health care. Such consequences not only contradict the limited goals of state intervention but also pave the way for the acquisition of new and broader skills on the part of the state.

The common expectation was that the intensity of services in hospitals—that is, the quantity and variety of services within a given length of stay—would be reduced in order to maximize the DRG payments relative to resource use. The actual picture is more mixed. Some studies do report fewer services for Medicare inpatients, indicating that

---

[14] Other indicators of hospital decline that began with the onset of the DRG system are: decreases in the number of persons employed by hospitals, growth in capital costs, decreases in revenue in relation to costs, and increases in hospital closings (PPAC 1988, chap 2).

the DRGs have had a negative impact on intensity of treatment (Fitzgerald et al. 1987). However, other studies note that, while the total volume of services devoted to each patient is declining, the development is not out of line with past trends (Notman et al. 1987). Still other studies report that services have not decreased, indeed some have expanded. For instance, referrals for such ancillary services as physical therapy either have not changed as a result of the DRGs (Thienhaus and Simon 1987) or have increased significantly after the implementation of the DRGs (Dore 1987). And certain ancillary programs, such as those for infection control, are continuing to develop or are actually being newly initiated (Haley et al. 1987). Physicians are also reporting that they are "condensing" services into the shorter period of hospitalization, that is, maintaining or increasing the quantity or variety of treatments (Notman et al. 1987, pp. 1263–67). Finally, the use of medical devices and the number of surgeries performed both actually increased between 1984 and 1986 (PPAC 1988, pp. 21–25).

It appears, then, that although the incentive structure of the DRG system might be seen as encouraging fewer services in order to reduce costs, a twisted scenario is occurring. Government intervention to hold down costs is in some cases precipitating more medical intervention. The increase in the number of surgeries, which are among the higher yielding DRGs, is particularly pertinent.[15] In addition, an increased volume of care has occurred as hospitals concentrate on certain specialized procedures, mostly involving surgery, that are revenue generating (PPAC 1988, p. 23). Thus, not only has cost containment not yet led to the *kind* of reductions in services that had been expected, it is also fashioning a new system of health care delivery that is distorting the minimalist rationale of price-setting guidelines into a more maximalist restructuring.

Instead of reductions in services, alternative methods of enhancing efficiency and/or income in delivering care are evolving. Most notable in this regard is the shift in where and how established treatments are being conducted. In both location and mode of treatment, out-of-hospital settings are being used much more than in the past. For example, there is a pronounced increase in the number of procedures being performed on an outpatient basis, including both minor surgical and nonsurgical, such as bronchoscopies and chemotherapy (Notman et al. 1987, pp. 1263–64). By 1988 the PPAC (p. 11) concluded that "substituting outpa-

---

[15] It should be noted that surgeries are among the higher yielding DRGs because the DRGs were originally calculated on an average past cost basis. Under the alternative criteria of the RBRVS scale discussed below, reimbursement to surgeons under Part B of Medicare will be lower than in the past

tient care for part of an inpatient stay or for an entire admission is one of the most important changes in practice patterns of the 1980s."

Out-of-hospital care is a response to the pressure to minimize hospital stays and episodes, a response that seemingly advances the interests of providers because outpatient services are not subject to the DRG system. Indeed, because in the outpatient sector physicians are reimbursed on the original fee-for-service basis and hospitals on the original cost-based system, the government could find that controlling one area of expenditure drives Medicare costs up elsewhere. (A potential offsetting measure, the resource-based relative value scale, is discussed below.) A more profound implication of the use of outpatient facilities, however, is that by thus "working the system" providers may actually be contributing to the long-range success of the structural changes suggested by regulation. New information on severity-of-illness cases suggests that providers are learning to use both inhospital and outpatient facilities more efficiently and effectively. Whereas hospitals were once used routinely for all kinds of cases, now more care than previously thought possible is being provided out-of-hospital. And the patients who are admitted to and treated in hospitals are, increasingly, the more severely ill, not those who can be treated quickly and discharged early (Director, HCFA, personal interview, August 1988; PPAC 1988, pp. 11–12). In short, it appears that in responding to economic incentives, health care providers are not only reconsidering previous conceptions about the locus of care but also reaffirming values about the need for care. The exercise is bringing previously idiosyncratic variables in medical practices and decisions into a more systematic management of key health care issues (Feinglass 1987). A major rationalization in American health care may yet result from government's single attempt to make hospitals more efficient.

A survey of physician perceptions of hospital pressures and incentives to change their practices concludes that "Medicare's PPS has had a widespread impact on physician inpatient activity" (as reported in DHHS/HCFA 1985, chap. 6, p. 12). Physicians have long been implicated in the high costs of health care (McMahon and Newbold 1986). As I have shown, they are now beginning to harness their treatment modalities and are responding as expected to accommodate the average cost constraints represented in the DRG weightings. The DRGs are framing the decision-making process. Although the state has not directly interfered by dictating the specifics of change, leaving physicians to fill in the actual compliance details, the state has provided a tight framework within which change has had to occur. The result is that the macrostructure of the DRGs is leading to a micro-level restructuring in health care. And the change has been state led.

The DRG system successfully brought Medicare inpatient hospital costs under control during its first few years of operation, with the result that the rate of increase in *total* Medicare payments slowed from a high of 20.9% in 1980 to 8.5% in 1987 (PPAC 1988, p. 93). Moreover, the 1990 expenditures from the federal Hospital Insurance Trust Fund, which pays hospital bills for Medicare beneficiaries, were about "$12 billion less in 1980 dollars, and $18 billion less in 1990 dollars, than was expected shortly before prospective payment went into effect—the equivalent of a savings of approximately 20 percent" (Russell and Manning 1989, p. 439). It is estimated that one-third of this saving is occurring because of the consequences of implementing the DRG system—above all because of the decline in hospital admissions, rather than from the prospective rates themselves (Russell 1989, p. 72). Since the late 1980s, however, hospital costs per case have been rising (PPAC 1990, pp. 25–26). Nevertheless, the exercise of successfully setting rates is having an impact beyond the sphere of hospitals alone. For the state has acquired a newfound strength and is now prepared to extend its reins of control to a formerly forbidden field, physician payment, in order to check still further the unintended consequences of fragmented policy mechanisms.

Physician Payment: The Next Step in Price Setting

Physicians have long cherished the ideal that they ought to "do all they can for their patients without taking into account the kinds of factors . . . that policy makers rightly should consider" (Morreim 1985, p. 31). From the available evidence it seems that physicians are now being constrained increasingly to consider those very policy factors. Policymakers have formulated a macrosystem of constraints that are both clinical and economic and have left it to hospital administrators and physicians to find methods of accommodation. The task of accommodation is preparing physicians for their more direct involvement in the framework of a changing system of health care, for the success of hospital-based DRGs in controlling Medicare expenditures is laying the foundation for yet further intervention.

As a system of prepayment, DRGs are applicable only to hospitals and affect physicians only in their capacities as hospital employees or users of hospital facilities. In 1985, however, Congress instructed that the feasibility of developing a prospective payment system for physicians also be investigated. It established a special task force, the Physician Payment Review Commission, to supervise studies and advise on findings. In 1988 the commission extended its full approval to a method of setting physicians' fees that had been under research supported by HCFA for several years. The resource-based relative value scale (RBRVS) measures various

dimensions of physicians' time and input in several services, procedures, and specialities (Hsiao 1987; Hsiao et al. 1988). The relative value scale, as a codification of standard procedures and services and as a measurement of resource costs for physicians, is not yet a revolutionary departure from past insurance practice. Blue Cross and Blue Shield have been using a type of relative value scale, based on physician charges, for over two decades. And the newly proposed system uses the same list of over 7,000 procedures and services developed by physicians to bill Blue Cross and Blue Shield, but it assigns resource-based values to the categories. Some physicians, especially surgeons who stand to lose under the RBRVS, oppose it. However, the American Medical Association considers the new scale an improvement over the "customary, prevailing, and reasonable" charge system, and has been involved in developing the resource-based values (AMA 1988, p. 3). Other groups, such as the American Association of Retired Persons, the Family Medicine Association, and the General Internists in Medicine, have also voiced their endorsement.

Like the DRGs, the relative value system represents a major instrument of change. The proposal embodies more than a benign system of standards for reimbursing physicians' fees: it represents a role for government in shaping the system. The new regulatory mechanism, which is now intended only to reflect prevailing norms and prices, could potentially intrude deeply into the pricing mechanism and the autonomy of private actors. It may be beginning as inauspiciously as the DRGs began, but the RBRVS may end up as a similar system of constraint and management.

The evolution of the DRGs to date has shown a steady progression in the readiness of government to intervene in correcting provider attempts to circumvent economic policy goals in health care. The growth of the DRGs also demonstrates the government's increasing role in developing a payment process that adheres to ever stricter limits on governmental as well as overall expenditures on health care. Government is no longer inclined to accept charges set by health care providers but is instead prepared to lead the pattern of funding in its own direction. In turning now to the last stronghold of medical prerogatives—physicians' fees— the government brings with it certain lessons learned from the experience of implementing the DRGs. As the fee schedule agenda unfolds, we might expect to see a replay of the major shifts that have occurred or may yet occur in the role of government in regulating the DRGs—above all, a shift from consensus-based pricing of inputs to more unilateral management of prices by the state. In liberal democracies change occurs incrementally. But as the discussion here has demonstrated, each case of limited state intervention sets the stage for further and more intrusive

intervention. This paper has identified a dynamic in state intervention in several areas of health policy. There is no reason we should not expect a repeat of the cycle in another area of health policy.

CONCLUSION: BEYOND LIBERAL INTERVENTION

In the United States we are witnessing a transformation in the nature of state intervention in which the central tenet of a liberal welfare state, namely the separation of the domains of state and society, is being transgressed. There are at least two dimensions of this transgression. One subordinates the principle that "the welfare state is a limited adjunct to the market, serving to correct the outcomes it generates but not seeking to displace it" (Moon 1988, p. 5). In the past, state intervention was aimed at the *recipients* of care, and at the consequences of market forces on the needy. Now it is aimed at the standard-bearers of the market in health care: the *owners* of facilities (hospitals) and, especially with the introduction of the relative value scale, the *producers* of services (physicians). A second transgression violates the principle that "the state should not be the architect of social order" (Anderson 1984, p. 15). While the DRG system contains no intended micro-level directives beyond implementing 475 categories of illness, it has nevertheless resulted in microstructural changes. The notion that government is getting out of the business of deciding on issues of medical care is belied by a host of new practices, from negotiations over categories to recalibrations of weightings and prices. And the consequence of this new form of state intervention has been the initiation of a new system of health care, exhibiting a publicly managed method of allocation, which in its own peculiar fashion represents a functional equivalent to the administration of nationalized health care that exists in some other welfare states.

   Indeed, the trajectory of interventionism and its paradoxical consequences that have been discussed here may be leading toward a more unified, buyer-determined, and regulated system of payment for health care, thereby ushering in the development of an American style national health insurance system. The changes introduced thus far have not yet resulted in a health care system that satisfies critics of the American welfare state. But they have initiated the crucial process of restructuring the foundations of the health care system, a process that could lead to the only kind of "national" insurance program that the United States is capable of achieving. We must not expect a monolithic national health care system like the one used in Great Britain, but we can expect incremental steps inching us closer to a *federated* system of multiple insurance bases similar to but more diverse than the "national" insurance program used in Canada. The Canadian health insurance system developed from

different foundations than the U.S. system (Taylor 1978; Evans and Stoddart 1986). Nevertheless, Canada remains a similarly rooted liberal welfare state, and it has undergone similar cycles of increasing state intervention to sharpen specific features of the health care payment system.[16] Most important, the state in Canada has gradually acquired the capacity not only to coordinate the system of health insurance, but also to *lead* developments toward national goals for health care. Researchers and theorists of the welfare state can learn from the trial-and-error experiences of Canada.

The U.S. government is beginning to determine the direction of change in health care provision, a development that is exemplified in the implementation of the DRGs as well as the RBRVS. The government has discovered that prospective payment works, and is attempting to extend its reach beyond inpatient hospital care. Some states have adopted prospective payment methods for programs other than Medicare. And a number of insurance companies, both private and nonprofit, have either adopted a prospective payment system, or are experimenting with alternative forms of prospective payment, or are exploring the possibility of so doing (Guterman et al. 1988, pp. 72–73). What is critical for the development of national health insurance is less the particularities of prospective payment than the establishment of a government-supervised common framework for payment within which the quest for greater equality of access and provision can be undertaken. The first step in establishing a common framework is the ability of the state to regulate effectively the unanticipated and adverse consequences of private and market forces. The United States now has taken this first step.

Although it is an important first step for the continuing evolution of the American welfare state, it is still a small one, considering the extensive health care needs still to be met in the United States. I have not addressed the issue of expansion in health insurance toward adequate coverage for the vast numbers of uninsured. That is a second and more problematic step in the trajectory of welfare state development. But I

---

[16] For example, the introduction of a public insurance program was initially acceptable to Canadian physicians because their fee-for-service incomes were maintained if not enhanced through the reimbursement system as well as through the possibility of charging patients an extra fee above the government's schedule of payments. As systems become institutionalized, however, changes develop. Eventually, the government in Canada prohibited the practice of extra billing, and provinces began to cap the annual increases in the fee schedules. Medical associations and provincial governments may negotiate on financial matters in reimbursement, but there is little doubt who has final authority in directing the relationship between health expenditures and the economy in Canada. Ironically, some argue that the degree of interference in physicians' practices is greater in the present American system than it is in Canada (Evans et al. 1989) or even Great Britain (Mechanic 1986, p 41).

have suggested that a necessary prerequisite—an institutional restructuring in the role of the state vis à vis the private sector—is beginning. This change in the role of the state lays the foundation, at least, for more equitable coverage and access in the future.

REFERENCES

Alford, Robert R. 1972 "The Political Economy of Health Care: Dynamics without Change." *Politics and Society* 2 (Winter): 127–64.

Amenta, Edwin, and Theda Skocpol. 1989. "Taking Exception: Explaining the Distinctiveness of American Public Policies in the Last Century " In *The Comparative History of Public Policies,* edited by Francis G. Castles. London. Polity.

American Medical Association (AMA). 1988. "The AMA and the Medicare Physician Payment System." Mimeographed, August.

Anderson, Charles. 1984. "Political Philosophy, Practical Reason, and Political Analysis." Paper presented to the American Political Science Association, Washington, D.C.

Brown, Lawrence D. 1986. "Introduction to a Decade of Transition " *Journal of Health Politics, Policy and Law* 11 (4). 569–83

Brown, Michael K., ed. 1988. *Remaking the Welfare State. Retrenchment and Social Policy in America and Europe.* Philadelphia: Temple University Press

Cromwell, Jerry, Sylvia Hurdle, and Rachel Schurman. 1987. "Defederalizing Medicaid· Fair to the Poor, Fair to Taxpayers?" *Journal of Health Politics, Policy and Law* 12 (1): 1–34.

Department of Health and Human Services/Health Care Financing Administration (DHHS/HCFA). 1985. *Report to Congress: Impact of the Medicare Hospital Prospective Payment System.* Baltimore: HCFA

Dore, Daniel. 1987. "Effect of the Medicare Prospective Payment System on the Utilization of Physical Therapy." *Physical Therapy* 67 (6): 964–66.

Esping-Andersen, Gòsta. 1989. "The Three Political Economies of the Welfare State." *Canadian Review of Sociology and Anthropology* 26 (1): 10–36.

Evans, Robert G., and Greg L. Stoddart, eds. 1986 *Medicare at Maturity: Achievements, Lessons and Challenges.* Calgary: University of Calgary Press.

Evans, Robert G., et al. 1989. "Controlling Health Care Expenditures—the Canadian Reality." *New England Journal of Medicine* 320.571–77.

Feinglass, Joe. 1987. "Variations in Physican Practice and Covert Rationing." *Theoretical Medicine* 8 (1)· 31–45.

Fetter, Robert B., et al. 1980. "Case Mix Definition by Diagnostic Related Groups." *Medical Care* 18 (2) suppl.. 1–53

Fitzgerald, John F., et al. 1987. "Changing Patterns of Hip Fracture Care before and after Implementation of the Prospective Payment System." *Journal of the American Medical Association* 258 (July 10): 218–21.

Guterman, Stuart, and Allen Dobson. 1986 "Impact of the Medicare Prospective Payment System for Hospitals." *Health Care Financing Review* 7 (3). 97–114.

Guterman, Stuart, et al 1988. "The First 3 Years of Medicare Prospective Payment: An Overview." *Health Care Financing Review* 9 (3): 67–77.

Gutmann, Amy, ed. 1988. *Democracy and the Welfare State.* Princeton, N.J.: Princeton University Press.

Haley, Robert W , John W. White, David H. Culver, and James M. Hughes. 1987 "The Financial Incentive for Hospitals to Prevent Nosocomial Infections under the Prospective Payment System." *Journal of the American Medical Association* 257 (March 27). 1611–14.

Hsiao, William C 1987. "The Resource-Based Relative Value Scale: An Option for Physician Payment." *Inquiry* 24 (4). 360–61.

Hsiao, William C., Peter Braun, Douwe Yntema, and Edmund R Becker. 1988. "Estimating Physicians' Work for a Resource-Based Relative-Value Scale." *New England Journal of Medicine* 319 (September 29): 835–41.

Iezzoni, Lisa, and Mark A. Moskowitz. 1986. "Clinical Overlap among Medical Diagnosis-Related Groups." *Journal of the American Medical Association* 255 (February 21)· 927–29.

Iglehart, John K. 1986. "Health Policy Report: Early American Experience with Prospective Payment of Hospitals " *New England Journal of Medicine* 314 (May 2): 1460–64.

Kudrle, Robert T., and Theodore R. Marmor. 1981. "The Development of Welfare States in North America " Pp 81–121 in *The Development of Welfare States in Europe and America,* edited by Peter Flora and Arnold J. Heidenheimer New Brunswick, N.J.: Transaction.

Lagoe, Ronald J. 1986. "Differences in Hospital Discharge Rates: A Community-based Analysis." *Medical Care* 24 (9): 868–72

Light, Donald W 1991. "The Restructuring of American Health Care." Pp. 53–65 in *Health Politics and Policy,* 2d ed. Edited by Theodor J. Litman and Leonard S Robins. Albany, N.Y.. Delmar

Light, Donald W., and Sol Levine. 1989. "The Changing Character of the Medical Profession." *Milbank Memorial Fund Quarterly* 66 (1): 1–23

McMahon, Laurence F., and Robert Newbold. 1986 "Variation in Resource Use within Diagnosis-related Groups." *Medical Care* 24 (5): 388–97.

Mechanic, David. 1986. *From Advocacy to Allocation: The Evolving American Health Care System.* New York: Free Press.

Melloan, George. 1988. "How to Save $30 Billion on Defense, in Theory." *Wall Street Journal,* September 20

Moon, J. Donald. 1988. "Introduction· Responsibility, Rights, and Welfare." Pp. 1–15 in *Responsibility, Rights, and Welfare: The Theory of the Welfare State,* edited by J. Donald Moon. Boulder, Colo. Westview.

Morone, James A., and Andrew B. Dunham. 1985. "Slouching toward National Health Insurance: The New Health Care Politics." *Yale Journal on Regulation* 2 (2): 263–91.

Morreim, E. Haavi. 1985 "The MD and the DRG." *Hastings Center Report,* June, pp. 30–38.

Notman, Mark, et al. 1987. "Social Policy and Professional Self-Interest: Physician Responses to DRGs." *Social Science Medicine* 25 (12). 1259–67

Perloff, Janet D , Phillip R Kletke, and Kathryn M. Neckerman 1987 "Physicians' Decisions to Limit Medicaid Participation: Determinants and Policy Implications." *Journal of Health Politics, Policy and Law* 12 (2): 221–35.

Piven, Frances Fox, and Richard A. Cloward. 1971. *Regulating the Poor: The Function of Public Welfare.* New York: Vintage.

Prospective Payment Assessment Commission (PPAC). 1988. *Medicare Prospective Payment and the American Health Care System: Report to the Congress.* Washington, D C.: PPAC, June.

———. 1990. Report and Recommendations to the Secretary, U.S Department of Health and Human Services. Washington, D.C.: PPAC, March 1.

Quadagno, Jill. 1987. "Theories of the Welfare State " *Annual Review of Sociology* 13:109–28.

Robinson, James C., Harold S Luft, Stephen J. McPhee, and Sandra S. Hunt. 1988. "Hospital Competition and Surgical Length of Stay " *Journal of the American Medical Association* 259 (February 5): 696–700.

Rothschild, Emma 1988. "The Real Reagan Economy." *New York Review of Books,* June 30, pp 46–53.

Ruggie, Mary 1984. *The State and Working Women. A Comparative Study of Britain and Sweden* Princeton, N.J.: Princeton University Press

———. 1990. "Retrenchment or Realignment? U.S. Mental Health Policy and the DRGs." *The Journal of Health Politics, Policy and Law* 15 (1). 145–67.

Russell, Louise B. 1989. *Medicare's New Hospital Payment System: Is It Working?* Washington, D.C. Brookings.

Russell, Louise B., and Carrie Lynn Manning. 1989. "The Effect of Prospective Payment on Medicare Expenditures." *New England Journal of Medicine* 320 (February 16). 439–44.

Schieber, George J , and Jean-Pierre Poullier. 1991. "International Health Spending: Issues and Trends." *Health Affairs* 10 (1) 106–16.

Skocpol, Theda. 1988. "A Society without a 'State'? Political Organization, Social Conflict, and Welfare Provision in the United States " *Journal of Public Policy* 7 (4) 349–71.

Smith, Steven Rathgeb, and Deborah A Stone. 1988. "The Unexpected Consequences of Privatization " Pp. 232–52 in *Remaking the Welfare State: Retrenchment and Social Policy in America and Europe,* edited by Michael K. Brown Philadelphia Temple University Press

Spicker, Stuart F., ed. 1987. "Prospective Payment: DRGs and Ethics." *Journal of Medicine and Philosophy* 12 (2)

Starr, Paul. 1982. *The Social Transformation of American Medicine.* New York: Basic.

———. 1989. "The Meaning of Privatization " Pp. 15–48 in *Privatization and the Welfare State,* edited by Sheila B. Kamerman and Alfred J. Kahn Princeton, N J.: Princeton University Press.

Stevens, Robert, and Rosemary Stevens 1974. *Welfare Medicine in America: A Case Study of Medicaid.* New York. Free Press.

Taylor, Malcolm G. 1978. *Health Insurance and Canadian Public Policy. The Seven Decisions that Created the Canadian Health Insurance System.* Montreal: McGill-Queen's University Press

Thienhaus, Ole J., and Sandra E. Simon. 1987 "Prospective Payment and Hospital Psychiatry." *Hospital and Community Psychiatry* 38 (10). 1041–43.

Thompson, Frank J. 1987. "Review Essay. The Politics of DRGs." *Journal of Health Politics, Policy and Law* 12 (4) 717–24

Thurow, Lester C 1985. "Medicine versus Economics." *New England Journal of Medicine* 313 (September 5): 611–14.

U.S. Congress 1966. *United States Statutes at Large,* 89th Cong., 1st sess., vol. 79 Washington, D.C : Government Printing Office.

Vladeck, Bruce C. 1979. "The Design of Failure: Health Policy and the Structure of Federalism " *Journal of Health Politics, Policy and Law* 4 (3). 522–35.

——— 1980. *Unloving Care: The Nursing Home Tragedy* New York: Basic.

Weir, Margaret, Ann Shola Orloff, and Theda Skocpol, eds. 1988. *The Politics of Social Policy in the United States* Princeton, N J.: Princeton University Press.

Wilensky, Harold L. 1975 *The Welfare State and Equality.* Berkeley: University of California Press

Sociological and Economic Theories of Markets and Nonprofits: Evidence from Home Health Organizations[1]

Lee Clarke
Rutgers University

Carroll L. Estes
University of California, San Francisco

In this article, data on home health agencies reveal surprisingly few differences between organizations of different tax status (nonprofit vs. forprofit). These findings directly contradict economic theories of nonprofit organization, as well as third-party government theory from political science. Since ecological theory in sociology proposes a useful explanation for organizational similarities, the authors search for differences between nonprofit and forprofit organizations among generalists and specialists. They find that specialists are quite similar, regardless of tax status, but generalists are different, a disparity that seems to contradict ecological analysis. Since institutional theory also explains isomorphism, the authors search for differences between organizations that do and do not belong to chains. Chains are more similar than nonchains. The discussion of the results centers on changes in the institutional context of home health care over the last decade, especially cascade effects of prospective payment schemes and the advent of new organizational forms. The authors conclude that the idea of accommodative isomorphism helps account for similarities among nonprofit and forprofit chains.

How should we theorize nonprofit organizations in the political economy? Most work on nonprofits, especially in economics, assumes legal form matters (Powell and Friedkin 1987; Steinberg 1987; Estes and Alford

1990). Nonprofit organizations' status in the Internal Revenue Service code (501(c)(3)) confers tax relief and "proscribes distributing net income as dividends or above-market remuneration" (DiMaggio and Anheier 1990, p. 138). This nondistribution constraint, as it is termed, is the theoretic cornerstone of economic theories of nonprofit organizations. As Hansmann (1987, p. 28) notes, economic "theories of the nonprofit firm are . . . essentially theories of the way in which the presence of a nondistribution constraint affects a firm's role or behavior."

Two complementary economic theories are particularly important. The first, deriving from the so-called principal-agent problem, is contract-failure theory (Hansmann 1987). According to this view, under some conditions consumers are particularly disadvantaged vis-à-vis service providers regarding information on quality and quantity of the service or product they desire. This is a contract failure because of the high "cost of developing and enforcing a contract that specifies" hard-to-measure product attributes (Weisbrod 1989, p. 145). For example, it is difficult for patients to judge either quantity or quality of health care services. Health care patients also often have limited mobility, so they are unlikely to comparison shop for services. "In such circumstances," says Hansmann (1987, p. 29), "a forprofit firm has both the incentive and the opportunity to take advantage of customers by providing less service to them than was promised and paid for."

Moreover, forprofits can more easily meet excess demand by raising prices, while nonprofits are constrained by a broader array of external forces such as public service mandates and the expectations of professionals, communities, politicians, and boards of directors. Broader mandates, in turn, lead nonprofits to use waiting lists as mechanisms to meet excess demand, as Weisbrod (1988) discovered was true for nursing homes and facilities for the mentally handicapped.

Fearful of forprofit opportunism, clients search for a mechanism that "mitigates the incentives of the agent [organization] to act contrary to the interests of the principal [client]. The nonprofit corporate form, with its nondistribution constraint, serves this purpose" (Hansmann 1987, p. 31; see also Shortell et al. 1986). In other words, rather than monitoring organizational output, clients rely on legal form as a signal that the urge to profit maximize has been muted and, hence, that nonprofits merit institutional trust (James 1987, 1989; Weisbrod 1988).

The second strand of economic theory—government-failure theory—is not inimical to the contract-failure hypothesis but develops its logic in a different manner. Burton Weisbrod (1988), for instance, agrees that nonprofit status symbolizes that quality will not be sacrificed for profits, but he argues further that "when demand is diverse . . . whatever quantities and qualities of services government provides will over-satisfy some people and under-satisfy others" (Weisbrod 1988, p. 25). In this view, government agencies tend to satisfy the average voter, resulting in excess demand for certain types of services. The chief function of nonprofit organizations is to fill the gaps left by this governmental failure. Government-failure theory holds that nonprofits and forprofits have distinct portfolios of services for different types of clients. Because contract- and government-failure theory are not contradictory explanations, we refer to them under the same rubric, "economic theory."

In this article, we contribute to the scholarly discussion by addressing the question: What are the differences between nonprofit and forprofit organizations in home health care? Contrary to much current thought on the matter, we find surprisingly few differences. We first review previous work that stresses differences between nonprofits and forprofits in health care. Next, using data on home health agencies, we demonstrate substantial similarity on a number of variables that were available. We draw on organizational, political, and economic theory to interpret our findings.

DIFFERENCES BETWEEN HEALTH CARE ORGANIZATIONS

A number of endogenous and exogenous dynamics are producing apparently momentous changes in the organization of U.S. health care. The foremost endogenous factor is the advent of new organizational forms, especially among hospitals (Ermann and Gabel 1984; Guterman and Dobson 1986; Guterman et al. 1988) but increasingly among other institutional health care providers as well (Light 1991). Industry structure is being reshaped by the expansion of chains (horizontal integration), mergers (Luke, Begun, and Pointer 1989), and multifacility systems (vertical integration) (Starr 1982, chap. 5; Levitz and Brooke 1985; Liszewski and Griffith 1988). While some believe nonprofits are increasingly seeking such forms (Friedman and Shortell 1988; Shortell 1988), traditionally they have been assumed to be the province of forprofit organizations.

Indeed another endogenous change in health care is the rise of forprofit organizations, although forprofits certainly are not new to medical service delivery (Trager 1990; Light 1991). Between 1972 and 1986, the number of Medicare-certified forprofit home health agencies grew from 42 to 1,869, an enormous growth rate (Hall and Sangl 1987; Gornick and Hall 1988; Marmor et al. 1987). Shamansky (1988, p. 387) notes that "the

provision of home care services used to be the purview of [nonprofit] health nursing agencies or visiting nurse associations, [whereas] today home care has become a multibillion dollar business, much of which is located in the forprofit arena."

The chief exogenous dynamic of change in health care concerns how Medicare reimburses hospitals. Congressional legislation (in 1983) shifted the basis of Medicare reimbursement to hospitals from *cost* of treatment (which allows cost to float with medical attention) to *prospective payment* (which ties cost to an expected average for a diagnosis related group) (U.S. Congress 1983). ("Prospective payment" is the generic term for the system; "diagnosis related group" or DRG, is the method used to implement prospective payment.) For example, under prospective payment a hospital will receive equivalent payments, ceteris paribus, for two colostomies, regardless of the length of patient stay. If one patient requires a hospital stay of 10 days, and the other requires four, but the DRG payment covers only four days, the hospital absorbs the extra six days. Under cost-based payment, all 14 days would have been reimbursed. If yet another patient is treated in fewer days than the DRG covers, the hospital keeps the difference.[2]

Prospective payment shifts the locus of fiscal incentive from physician-driven demand to conformity with bureaucratic categories (DRGs); it also gives hospitals an incentive to appropriate the fiscal difference between mandated reimbursements and shortened length of stays (Schlesinger and Dorwart 1984; Feder, Hadley, and Zuckerman 1987; Glazer 1988). One result, some evidence suggests, is the "quicker and sicker" syndrome, wherein patients are discharged sooner and sicker than was the case previously (Bergthold et al. 1988; Guterman et al. 1988; Neu and Harrison 1988; Schramm and Gabel 1988; Gay et al. 1989). Greater pressure to minimize costs also encourages accelerating the rate of patient discharge to cover "cost outliers."

As the foregoing suggests, hospitals have received the lions' share of scholarly attention (Pattison and Katz 1983; Relman 1983; Alexander, Morlock, and Gifford 1988; Friedman and Shortell 1988), leaving significant gaps in what we know about some types of service providers. Fortunately, there is a burgeoning literature on the larger set of organizations that provide health-related services, especially to the aged, including: home health agencies (Leader 1987; Olson and Olson 1988; Wood and Estes 1990), adult day- and health-care centers (Estes et al. 1988), hospices (Paradis and Cummings 1986), and nursing homes (Vladeck

---

[2] The DRGs are actually a more drastic change than this. Part of the intent behind creating this system was to encourage fiscal and medical innovation in hospitals. Detailed treatment of this issue is beyond this paper's scope.

1980; Harrington 1984; Neu and Harrison 1988; Swan, de la Torre, and Steinhart 1990). Such organizations are indirectly affected by the macro-level forces mentioned above. If, for example, hospitals discharge patients sooner and sicker, the demand for nursing home and home health services increases and shifts upward.[3]

As the typical hospital stay has shortened, home health organizations increasingly provide services once provided in hospitals (Bergthold et al. 1988; Bergthold, Estes, and Villanueva 1990; Spohn, Bergthold, and Estes 1988). Newcomer, Wood, and Sankar's (1985) prediction that demands on home health organizations would increase because of shorter hospital stays was substantiated by Guterman et al. (1988) and Neu and Harrison (1988). Recent data from the Health Care Financing Administration show that about 1.5 million patients were served by Medicare-certified home health care agencies in 1989, compared to about 726,000 in 1980 (Prospective Payment Assessment Commission [PROPAC] 1990, p. 69). In 1989 there were 1,172 home health care visits per 1,000 Medicare enrollees, compared to 577 in 1980 (U.S. House of Representatives, Committee on Ways and Means 1990, p. 144). Home health care is clearly important in the health care industry, and is likely to become even more so in the future.

DATA AND MEASUREMENT

Data

This research is part of a broad effort by the Institute for Health and Aging at the University of California, San Francisco, to assess the consequences of federal policy changes regarding health care payments (Estes et al. 1988). Our random samples were drawn in 1987 in nine standard metropolitan statistical areas (SMSAs), as defined by the Census Bureau, in five states (California, Florida, Pennsylvania, Texas, and Washington).[4] The SMSAs served as the sampling frame because they roughly correspond to market areas for health care agencies.

Interviewers telephoned 198 home health agencies drawn from 1986 lists of providers held by state licensing and certification agencies, county lists where available, and trade associations. The study was designed to permit a sufficiently large sample of each provider type (with the excep-

---

[3] Other influences on demand for home health services include an aging population, more home-based technologies, and more Medicare claims from people under 65 with mental illnesses, physical disabilities, and AIDS for whom private reimbursement schemes are rare (Shamansky 1988).

[4] The SMSAs were San Francisco/Oakland, San Diego, Miami, Tampa/St. Petersburg, Philadelphia, Pittsburgh, Dallas/Ft. Worth, Houston, and Seattle.

tion, in some cases, of government agencies) to enhance representativeness and to support inferences to the universe of private agencies in an SMSA.

To be selected, a home health agency (HHA) had to provide at least skilled nursing care and home health aide services; at least 10% of an agency's clientele receiving these services had to be 65 years old or older.[5] The sample included hospital affiliated HHAs as well as independent, freestanding agencies; agencies were included whether or not they were Medicare certified. If an organization had a Medicare-certified branch *and* a noncertified branch at the same address, it was counted as a single agency. If we later discovered such organizations had separate budgets and program statistics, separate directors with no overall director located at the address, or were considered separate by the director, both were given an equal likelihood of selection with all other agencies. This procedure added 15 non-Medicare-certified HHAs to the sample. Using state, county, and trade association lists rather than telephone directories to derive the HHA universe probably biased the sample toward larger, more established, licensed and certified agencies. Also, HHA tax status is somewhat biased toward forprofit providers because there are more of them in the western and southern regions of the country.

The response rate for eligible HHAs was 89%. The sample had 98 forprofit, 75 private nonprofit, and 3 government agencies (N = 176). We dropped the three government organizations from the analyses.

Measurement

Table 1 summarizes our variables and their definitions.

Tax status.—Since scholarly (and policy) debates about U.S. health care have centered on differences between nonprofit and forprofit agencies, our primary independent variable is tax status. Of the 173 organizations, 57% were forprofits and 43% were nonprofit.

Control variables.—We use two controls in the basic analysis—organizational size and whether an organization belongs to a chain of similar organizations. Chain membership could confound relationships between tax status and dependent variables because chains generally have well-developed mechanisms for securing hospital referrals, a major conduit for home health client acquisition. Chains are also more prevalent among forprofits than nonprofits (Waldo, Levit, and Lazenby 1986;

---

[5] This criterion was used because of an overall project focus on providers of services to the aged. The criterion resulted in excluding less than 3% of the universe of home health organizations from the sample frame Thus it is unlikely that the criterion negatively affects the generalizability of our findings

TABLE 1

VARIABLES AND DEFINITIONS

Variables	Definitions and Coding
Independent and control variables·	
Tax status	Private nonprofit (coded 0); forprofit (coded 1)
Chain	Belongs to a parent corporation operating other home health agencies (coded 1, 0 otherwise)
Size	Total number of staff on the agency's payroll
Dependent variables:	
Medicare certification	Any part of the organization is Medicare certified (coded 1; 0 otherwise)
Service provision·	
No. of services	Total no. of services provided
Infusion therapy	Provides (coded 1) or does not (coded 0)
Hospice services·	Provides (coded 1) or does not (coded 0)
Physical therapy	Provides (coded 1) or does not (coded 0)
Aide services	Provides (coded 1) or does not (coded 0)
Homemaker	Provides (coded 1) or does not (coded 0)
Ventilator	Provides (coded 1) or does not (coded 0)
Social work	Provides (coded 1) or does not (coded 0)
Waiting list	Maintains list (coded 1) or does not (coded 0)
Client mix:	
65 and older	% of clients aged 65 and older
Low income	% of clients low income
Minority	% of clients minorities
Clients served each month	Total no. of clients served per month
Contracting.	
Acute care hospitals	Contracts with (coded 1) or does not (coded 0)
Psychiatric hospitals	Contracts with (coded 1) or does not (coded 0)
Nursing homes	Contracts with (coded 1) or does not (coded 0)
Board and care	Contracts with (coded 1) or does not (coded 0)
Other home health agencies	Contracts with (coded 1) or does not (coded 0)
Senior centers	Contracts with (coded 1) or does not (coded 0)
State health department	Contracts with (coded 1) or does not (coded 0)
Hospices	Contracts with (coded 1) or does not (coded 0)
Health maintenance organizations	Contracts with (coded 1) or does not (coded 0)
Preferred provider organizations	Contracts with (coded 1) or does not (coded 0)
Staffing:	
Aides	Total no of home health aides
RNs	Total no. of registered nurses
LVNs/LPNs	Total no. of licensed vocational/practical nurses
PTs	Total no of physical therapists
Revenue sources	
Medicare	% revenue from Medicare
Medicaid	% revenue from Medicaid
Insurance	% revenue from insurance
Copayments	% revenue from copayments
Philanthropy	% revenue from philanthropy
Government dependence	% revenue from Medicare and Medicaid
Generalism	Generalists (coded 1) if number of services provided was greater than mean (5.25), specialists coded 0
Herfindahl index	Measures market concentration; the closer to 1, the less competitive the market.
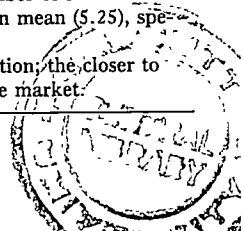
Bergthold et al. 1988), perhaps impeding nonprofits' ability to compete with forprofits (Salvatore 1981, 1982). Fifty-four percent of our forprofit HHAs belonged to chains, compared with 9% of the nonprofits. Organizations were coded positively on chain membership if they belonged to a parent organization operating other home health agencies. Competition was estimated with an approximated Herfindahl index (for more on this, see n. 13 below).

We measured agency size by the total staff an agency employed. The range for size was 3 to 617, with a mean of 92. (The Appendix contains means, SDs, and sample sizes for all variables.)

Dependent variables.—The dependent variables tap most key attributes of home health agencies.[6] Medicare certification, measured by asking if organizations were so certified, must precede receiving Medicare money, a main revenue source. Significantly more nonprofits were Medicare certified (97% vs. 87% of the forprofits), though these percentages suggest little substantive difference. The service provision group taps all the major services home health agencies provide. We also asked about waiting lists. It is interesting that the vast majority of forprofits (94%) and nonprofits (92%) maintained no such list.

Client mix indicates the percentage of an organization's total clientele represented by each group. These variables are interesting because both economists and political analysts argue that nonprofits will more likely serve the old, the poor, and (because of their disproportionate numbers among the poor) minorities. Hollingsworth and Hollingsworth (1989), for example, found such differences in a study of hospitals in England and Wales. We also measured how many clients an HHA served per month.

The contracting variables represent the types of organizations home health agencies contracted with. Home health agencies use two general classes of contracts, both for labor: daily contracts with individual professionals ("per diem" contracts) and nursing registries. Registries cost more than daily contracts, because the contractor assumes some of the profit, which is one reason per diem contracts are more widely used. Contracting with individuals lets home health agencies better control staff and overhead costs, because the workers receive no benefits and are not paid for travel time to and from patients' homes (which also saves on automobile insurance and maintenance).

Staffing patterns are important because home health is a labor intensive enterprise (Szasz 1990). Use of each employee type is reflected in how many of each category an agency employed. Finally, reliance on various revenue sources is measured by the proportion of an organiza-

---

[6] We lack data on service quality, a major gap in research on HHAs in particular and health care organizations in general (cf. DesHarnais et al. 1987).

tion's budget from Medicare, Medicaid, insurance, copayments (i.e., from clients, family members), or philanthropy.

COMPARING HOME HEALTH AGENCIES

Legal Form

We first ask: Does legal form matter?[7] Economic theory and medical policy analysis claim it does, both predicting nonprofits will serve more poor and minority clients, and, therefore, that nonprofits will depend more on Medicaid. The same logic forecasts forprofits will serve fewer low income clients and, since those clients are more likely to be underinsured or uninsured, depend more on copayments and insurance. Further, as mentioned, forprofits can raise prices to allocate resources; while nonprofits presumably use waiting lists. Finally, because forprofits are (presumably) more driven to maximize profits and because contracting in home health is usually cheaper than providing services in-house, forprofits should more extensively contract for services compared with nonprofits. This should be especially true for labor-intensive services.

Table 2 presents the unstandardized coefficients from regressions of each of our dependent variables on tax status. As indicated by the lack of significance ($P \leq .05$) for most coefficients, our chief finding is one of remarkably few differences. Even without controls, column 1 shows that only 14 of 34 bivariate relationships were statistically significant. Once we controlled for size and chain membership (col. 3), the number of significant coefficients dropped to six.[8]

The blend of nonsignificant and significant relationships in column 3 of table 2 has mixed implications for theories of nonprofit organization. Clearly, a set of nonfindings directly contradicts prevailing economic theory. For six of the seven types of major services, tax status does not meaningfully distinguish between organizations. For those six services, there are no systematic differences in the types of services nonprofits and forprofits provide. This is surprising, because we expected nonprofits to more likely provide services not usually covered by insurance (e.g., social work) and forprofits to cover insured services (e.g., ventilator care).[9] Also, nonprofits are not more likely to maintain waiting lists, contradicting predictions from economics. Only 12 of 173 agencies (6 of each organi-

---

[7] Weisbrod (1988) found large differences between secular and religious nonprofits. Indeed he found greater differences between religious and secular nonprofits than between secular nonprofits and forprofits. Unfortunately, we have no data on religious orientation.

[8] Logging size made no substantial difference in the significance levels.

[9] Ventilators provide continuous mechanical breathing assistance.

TABLE 2
EFFECTS OF TAX STATUS ON 34 DEPENDENT VARIABLES

| | UNSTANDARDIZED COEFFICIENTS | | |
DEPENDENT VARIABLES	Tax Status[a] (1)	Tax Status Net Chain (2)	Tax Status Net Chain, Size (3)
Medicare certified	-.106* (.973*)	-.080 (.979*)	-.074 (.996*)
Service provision			
No. of services provided	.298 (5.08*)	.175 (5.05*)	.052 (4.85*)
Provides infusion therapy	.031 (.867*)	.012 (.863*)	-.007 (.856*)
Provides hospice services	.003 (.467*)	-.024 (.461*)	-.046 (.440*)
Provides physical therapy	-.075 (.973*)	-.069 (.975*)	-.080 (.961*)
Provides aide services	-.007 (.987*)	.005 (.989*)	.003 (.984*)
Provides homemaker	.289† (.333*)	.206* (.316*)	.174* (.235*)
Provides ventilator	.180* (.514*)	.113 (.499*)	.067 (.407*)
Provides social work	-.144* (.960†)	-.086 (.972*)	-.078 (.985*)
Maintains waiting list	-.019 (.080*)	-.007 (.082*)	.003 (.064)
Client mix:			
% 65 and older	.381 (75.7*)	4.82 (76.6*)	5.16 (80.1*)
% low income	.627 (33.6*)	-2.71 (32.8*)	-1.37 (33.4*)
% minority	-2.74 (28.8*)	-1.47 (29.1*)	-1.66 (29.2*)
Clients served per month	-.185* (.317*)	-.174* (.319*)	-.189* (.235*)

Contracting patterns

Contract with acute hospitals	.021 (622*)	−.031 (.611*)	−.066 (556*)
Contract with psychiatric hospitals	−.051 (173*)	−.128* (.157*)	−.152* (.119*)
Contract with nursing homes	− 048 (.467*)	− 100 (456*)	−.123 (.385*)
Contract with board and care	.003 (.333*)	−.023 (.328*)	− 027 (.290*)
Contract with other home health organizations	−.013 (.533*)	−.081 (.519*)	−.083 (464*)
Contract with senior centers	−.096 (.333*)	−.116 (.329*)	−.106 (.285*)
Contract with state health department	−.019 (.213*)	−.067 (204*)	−.070 (.204*)
Contract with hospices	−.119 (507*)	−.161 (.498*)	−.190* (.437*)
Contract with health maintenance organizations	028 (.560*)	− 038 (546*)	−.080 (.485*)
Contract with preferred provider organizations	.124 (.284*)	075 (.273*)	.016 (.185*)

Staffing:[b]

No. of home health aides	26.3* (7.26)	21.4* (5.86)	7.93 (.630)
No of registered nurses	28.3* (11 5)	24 2* (10.5)	7.58 (3.61)
No. of LVNs or LPNs	42.2* (−8 30)	45 1* (−7.24)	1.11 (−1.23)
No of physical therapists	−.511 (4 51*)	− 171 (4.62*)	−.232 (4.17*)

Revenue sources:

% from Medicare	−20.5* (69.6*)	−10.9 (71.4*)	− 7.75 (80.8*)
% from Medicaid	−7.67* (5.35†)	−3.00* (5.37*)	−− 3.03* (6.83*)
% from insurance	10 8* (10 4*)	7.12+ (9 72*)	5 22 (5 58*)
% from copayments	18 3* (5.43)	12 2* (4 24)	8.87 (2 41)
% from philanthropy	−3 26* (3 49*)	−3 18* (3 50*)	−2 46* (2.39+)
Government dependence	−23 4* (75.3*)	−14.2* (76.9*)	−10.8 (86 6*)

NOTE —Constant terms are in parentheses We also ran logistic regressions for models with dichotomous dependent variables, none of the signs or coefficients changed
[a] Tax status is 0 = nonprofit and 1 = forprofit
[b] Controls for no of clients served per month to control for effects of staff/patient ratios
* $P \leq .05$

zation type) had waiting lists at all. Forprofit HHAs are not more likely than nonprofits to use waiting lists to allocate resources.

Nonprofits and forprofits also do not differ significantly on the proportions of low income and minority clients they serve. Nonprofits are somewhat more likely than forprofits to contract with psychiatric hospitals and hospices. The substantive meaning of these relationships is ambiguous, but prevailing theory would lead us to expect far more significant differences between nonprofits and forprofits in contract reliance. Nor are there significant differences between nonprofits and forprofits on any staffing variable. This is surprising because the health care literature suggests forprofit HHAs are more likely to hire staff skilled in services reimbursed by insurance, copayments, and Medicare. Finally, we find no differences between nonprofits and forprofits in percentage of income from Medicare, insurance, and copayments once we control for size and chain membership.

The startling paucity of differences between nonprofit and forprofit HHAs casts doubt on an economic explanation of nonprofit organizations. Recall that theory is founded on informational asymmetries between agents and principals (HHAs and clients); nonprofit tax status signals that organizations have fewer incentives to sacrifice service quality for profits (James 1989, p. 4). Since minority and low-income clients are more likely to be underinformed regarding health care services, we should find nonprofits serving such clients more than forprofits (Weisbrod 1988). Our data suggest that, at least for HHAs, clients are not using "nonprofitness" as a warrant to grant institutional trust. Of course, our findings do not mean poor and minority clients are, in fact, well informed about home health organizations. They do, however, contradict expectations given by available agent-principal theory.

A few significant differences are consonant with economic reasoning. Compared with nonprofits, forprofits are 17% more likely to provide homemaker services, net chain and organizational size. Marketing and regulation are two factors that help explain this difference. Forprofit agencies have aggressively marketed homemaker services, and, to the extent that forprofits are not Medicare certified, overhead costs for all services are lower than those of certified agencies because certified agencies must meet many costly regulatory and licensing requirements. So even if a certified, nonprofit agency offered homemaker services, those same types of services would be cheaper if delivered by a forprofit, uncertified agency.[10] When nonprofits offer homemaker services, their pro-

[10] We have limited support for this idea Of the 15 non-Medicare-certified agencies, 13 forprofits and 2 nonprofits provided homemaker services. The small number of noncertified organizations, however, makes this finding a footnote.

grams tend to be underwritten by non-Medicare federal funds. These funds are limited, so there may be a limit on the amount of homemaker services that nonprofits can afford to provide. In addition, homemaker services are generally excluded from Medicare reimbursement and are usually financed with copayments.

It is interesting that nonprofits serve considerably more clients per month than forprofits. This result likely reflects the tendency for non-profit HHAs to be older and more established community institutions, and such institutions serve more clients, in contrast to newer, less well established forprofit HHAs.[11] Forprofits also receive less of their revenue from Medicaid and philanthropy, although the coefficients are quite small.

OTHER POSSIBILITIES

Third-Party Government

Though economic theory affords only limited explanatory help, other aids are at hand. Salamon argues nonprofits are a "third-party government" wherein "the federal government performs a managerial function but leaves a substantial degree of discretion to its nongovernmental, or non-federal, partner" (Salamon 1987, p. 110). Rather than a response to government or market failure, as economic theory posits, nonprofits are a political success because the government can use them as subcontractors, permitting government expansion without incurring high political costs for doing so. Salamon (1987, p. 110) contends this system "creates a [democratic] public presence without creating a monstrous public bureau-cracy," though an equally plausible conclusion is that the nonprofit sector is a form of corporate welfare, channeling public monies to private pur-poses, a conclusion that is consonant with Estes and Alford's (1990) political economic theory.

Whatever the political conclusion, the theory of third-party govern-ment makes specific predictions about home health revenue sources. The theory might argue that forprofits are less dependent on government money because they are better positioned to serve paying clients and the insured. Conversely, nonprofits are more dependent because mandates force on them a broader, and poorer, clientele. Third-party government theory predicts significant differences for the degree of dependence on governmental sources of income.

To address this theory we created a variable that taps degree of depen-dence on government revenue sources, an additive function of Medicare

[11] Nonprofits were 15 years older, on average, controlling for chain and size.

and Medicaid. The results are shown in the last entry in table 2. The relationship between government dependence and tax status is significant, with nonprofits receiving 23% more of their budgets from government sources. Significance disappears, however, once chain and size are controlled. Although it is partially true that nonprofits depend more on government monies than forprofits, that dependency is attributable to size and chain membership rather than tax status.

Generalism and Specialism

To this point, an interesting difference in our data concerns the number of clients served per month, with nonprofits serving more clients per month than forprofits, even net chain membership and size. This finding supports the conclusion that legal form matters, but the lack of significance among so many of the other relationships suggests that tax status alone does not differentiate between HHAs. Perhaps the crucial distinction is whether HHAs are, in the language of ecological theory, generalists or specialists. DiMaggio and Anheier (1990, p. 144) say that "an ecological approach is well suited to test theories about the intersectoral division of labor." While we lack appropriate data for such a test, we can enlist some ideas from ecological analysis in our search for differences between nonprofit and forprofit organizations.

Specialists are usually small and provide fewer goods or services than generalists (Freeman and Hannan 1983; Hannan and Freeman 1984, 1989). Thus we might expect nonprofits to be generalists and to be larger and provide more services than forprofits. However, column 3 in table 2 shows that the relationship between tax status and number of services provided is insignificant. Also, although we do not present the findings in tabular form, including the number of services provided as a control variable did not further distinguish between HHAs of different tax status.

Still, perhaps tax status differences obtain *within* categories of generalists and specialists. To assess this possibility, we divided agencies into two groups according to number of services provided. No substantive reason dictated a particular cutoff, so we used mean number of services (5.25): organizations providing fewer than the mean were specialists ($n = 99$), those providing more than or equal to the mean were generalists ($n = 74$). We ran the models separately within categories of specialism and generalism to discover whether tax status matters within them. We list only the significant differences in column 1 of table 3.

First note that, contrary to expectations, forprofits are usually larger than nonprofits by about 45 employees (data not shown). However, we found a number of significant differences for tax status among specialists and generalists. Among specialists, nonprofits serve 105 more clients per

TABLE 3

SIGNIFICANT EFFECTS OF TAX STATUS WITHIN CATEGORIES OF GENERALISM/
SPECIALISM AND CHAIN/NONCHAIN

	UNSTANDARDIZED COEFFICIENTS	
DEPENDENT VARIABLES	Tax Status Net Chain and Size (1)	Tax Status Net Size (2)
Specialists (N = 92) [a]		
Clients served per month	−105* (189*)	
% from philanthropy ..	−1 95* (2.07*)	
Generalists (N = 63).		
Provides infusion services	−.107* (974*)	
Provides homemaker services	264* (590*)	
Clients served per month	−343* (344*)	
Contracts with psychiatric hospital	−.290* (.312*)	
% revenue from copayments	17.8* (2.79)	
% revenue from philanthropy	−3.47* (3 19*)	
Chain (N = 54)		
Contracts with hospices ...		−418* (.682*)
% revenue from Medicare		−40 8* (89.0*)
Nonchain (N = 101).		
Provides homemaker services		.210* (.194*)
Clients served per month		−220* (220*)
Contracts with psychiatric hospital		−.153* (.125*)
% revenue from Medicaid . .		−3.64* (5 97*)
% revenue from philanthropy		−2.78* (2 38*)

NOTE —Constant terms are in parentheses Tax status is 0 = nonprofit and 1 = forprofit.
[a] There were 99 specialists and 74 generalists, 60 agencies belonged to chains and 113 did not We present the smallest sample sizes for any equation within a category. The difference between 99 and 92, etc , is because of listwise deletion
* $P \leq .05$

month than forprofits and receive 2% more of their budgets from philanthropy. Of those differences only number of clients served per month is substantively nontrivial. Among generalists, more differences emerge. Nonprofit generalists are more likely to provide infusion services, while forprofits are more likely to provide homemaker services. Generalist nonprofits service 343 more clients per month than generalist forprofits and are more likely to contract with psychiatric hospitals. Concerning revenue, the differences on philanthropy remain substantively small, but generalist forprofits receive considerably more from copayments.

Specialist HHAs resemble one another more than generalist HHAs do, regardless of tax status. Specialists may tend to isomorphism because entry barriers (e.g., overhead, regulations) are lower for organizations willing to provide a more restricted array of services. Further, as our

sample distribution seems to show, there are more specialists in home health organizational fields, relative to generalists, which implies a greater degree of competition among specialists.[12] To assess this possibility we introduced a Herfindahl index, a rough measure of market share in an area, into the equations, but none of the significant relationships changed.[13] This suggests market competition is not, in fact, driving similarity among specialists. Moreover, why would specialists be alike but not generalists? Ecological reasoning suggests both generalists *and* specialists would be similar, because similar environments elicit similar responses from organizations competing for the same market niche (Hannan and Freeman 1977, pp. 940, 945; Carroll 1984).

Ecological analysis thus provides no clear explanation for more similarity among specialists as compared with generalists. An alternative explanation might be that generalist organizations tend to be older than specialists (by six years in our data) and possibly more established and protected from environmental pressures to conform. If so, future work should seek to discover whether generalists have more power than specialists in organizational fields. Another possibility is that groups and organizations that surround and are socially integrated with generalists have the requisite influence and resources to maintain distinctive organizations that serve their interests. These two explanations for similarity among generalists both center on power but locate it in different places. Either possibility would likely see as important social and political ties between organizations or perhaps between organizational elites. Examining those links may reveal asymmetrical patterns of resource exchange or elective affinities between ideological commitments and organizational form. Direct tests of power models of isomorphism are badly needed.

[12] We use the weak voice here not only because our data were not collected with the categories of generalism/specialism specifically in mind but, more important, because ours is a crude measure of degree of specialization and we do not have enough data on environmental characteristics.

[13] The Herfindahl index is given by

$$H = \sum_{i=1}^{N} Si^2;$$

where S_i is the market share of the ith organization (see Hirschman 1964, Scherer 1970; Swan and Estes 1990). The closer the index is to 1, the less competitive is the market. Our index was approximated by summing the squared relative proportion of average clients served per month by each HHA in an SMSA. We separately controlled for two other measures of competition, neither of which made any difference. (1) the proportion of HHAs per SMSA population (correlation between Herfindahl and agencies/population was −.53) and (2) the degree to which an SMSA varied from a 50/50 distribution of nonprofits and forprofits.

Chain Membership

There is a final possibility to explore. Our introduction noted the significance of new organizational forms such as chains, multifacility systems, and mergers. These developments suggest that the key classification may be chain membership and that tax status differences might emerge within those categories. So we divided the sample into chains and nonchains and reran the models. The second column of table 3 shows the significant differences in tax status among chains and nonchains, controlling for size. Among chains, forprofits are less likely to contract with hospices and are substantially less likely to receive Medicare revenues. On balance, home health agencies that belong to chains are isomorphic. More differences emerge for agencies that are not in chains. Among nonchains we find that forprofits are more likely to provide homemaker services and serve 220 fewer clients per month than do the nonchain nonprofits. Nonchain forprofits also are less likely to contract with psychiatric hospitals and receive a lower proportion of revenue from Medicaid and philanthropy.

The implications of either ecological or economic theory for these findings are not clear. That chains tend toward similarity is, however, consonant with an institutional analysis of organizational change A likely explanation for similarities among chains is a process of mutual organizational adjustment, a process that might be called accommodative isomorphism. Accommodative isomorphism occurs when there is movement of different types of organizations toward similar forms and behaviors, whether those forms and behaviors are extant in an organizational field or are new to it. This type of isomorphism suggests a process of organizational change that affects organizations irrespective of legal form. What is the source of such a process? One answer derives from the work of Meyer (1987) and his colleagues (Meyer, Stevenson, and Webster 1985), who have written cogently on the issue of bureaucratic growth.

Meyer et al. (1985) attribute a generic mechanism to organizations that drives the proliferation of bureaucratic units, personnel, and budgets. Organizations typically respond to persistent problems that defy existing categories (i.e., uncertainty) by creating more organization. Newly created units (offices, departments, etc.) try to resolve the original problems, but generate new problems of their own. They studied city finance departments in organizations that were typically larger than the organizations studied here; they also had measures of bureaucratization (e.g., differentiation) that we do not have. Thus the theoretic analogy we are proposing is not exact. What is useful about their argument here is the implication that, if organizations respond similarly to similar problems, then formal orientation toward profit is not the main concept to use for theorizing why organizations look and act similarly (or differently). From

this view, federal enactment of prospective payment certainly created new uncertainties—for example, regarding service delivery, personnel-use patterns, external relationships—for health care organizations regardless of their tax status. Before prospective payment, the chief constraint faced by forprofits was the demand of profit rates and fiscal responsibility; for nonprofits, more so than with forprofits at least, the chief constraint was public service. But prospective payment rearranged the resource mix available in the health care sector, increasing general pressure to respond to regulations and perhaps economic efficiency. And yet, given that we have not discovered uniform similarity among home health organizations, it seems likely that chains face different environmental influences than do nonchains.

Being a member of a chain exposes HHAs to similar forces, irrespective of tax status. Home health agencies in chains tend to be larger than nonchains, and there are fewer of them in any given organizational field. Thus it may be easier for chains to monitor each other's behavior, again regardless of formal orientation toward profit. Larger organizations that are relatively open to outside influences, in other words, are more likely than their counterparts to tend toward accommodative isomorphism.[14]

CONCLUSION

This study is the first we know of to use systematic data to assess differences between nonprofit and forprofit HHAs. However, because our data are synchronic we cannot assess unambiguously the mechanisms behind organizational similarities and differences in home health. Nevertheless, our data partially support some explanations of nonprofits and refute others. Still other findings require a new explanation.

Here we note some broad conclusions. First, legal form—tax status—does not usefully discriminate among HHAs. Hence theories based on an assumption of a legal-form difference risk founding their explanations on an insecure basis.[15] Moreover, on average, nonprofits and forprofits provide the same number of services, and this is also true within categories of chain membership and generalism/specialism. We thus find no support for the charge that forprofits serve better-paying, less troublesome clients

[14] This argument apparently contradicts our findings about specialists, i.e., while generalists and chains are larger than their opposites, there are more differences among the latter than the former This is because there is not a one-to-one correspondence between chains and generalists

[15] This conclusion parallels Perry and Rainey's (1988, see also Rainey et al. 1976) argument against assuming meaningful differences between public and private organizations.

while nonprofits are beneficent providers. Of course, we cannot tell if this means nonprofits act like forprofits or vice versa; still, the vast majority of our findings with respect to legal form are contrary to what economic theory predicts.

Second, once size and chain membership are controlled, nonprofits are not more dependent on government revenue than forprofits. Our analyses do not support the notion that nonprofits are a "third-party government."

Third, considering specialist and generalist HHAs, tax status is more useful for discriminating among generalists than specialists. While these findings seem to accord with ecological theory, at least for specialists, when we introduced our (admittedly limited) measures of competition the number of significant relationships did not change. Thus, ecological analysis does not help distinguish between HHAs of different tax status.

Finally, considering chain and nonchain HHAs, tax status is more useful for discriminating among nonchains than chains. These findings are not easily explained by either economics or ecology, although they are somewhat consonant with institutional theory.

Future explanations of organizational behavior in home health care should grant causal primacy to both the advent of prospective payment and a dramatically expanded presence of forprofits. Both forces changed the profile of available resources in organizational fields, shifting the demand for home health services upward, increasing environmental complexity by increasing the variety of organizational forms, and changing the nature of government reimbursement. It may be that these rearrangements compelled chains and specialists toward similar strategies and structures (Ritzer and Walczak 1988) to remain viable players for health care resources, while driving nonchains and generalists to differentiate themselves. However, the effects of these two macroforces have been indirect, and the precise mechanisms through which they operate await further research. Although home health agencies are not directly subject to prospective payment, it is likely that changes in Medicare reimbursement techniques have induced nonprofits to rationalize both internal practices (Marmor, Schlesinger, and Smithey 1987; Glazer 1988; Estes and Alford 1990) and outward appearances (Zucker 1988).

There has also been unmistakable expansion of forprofit organizational forms in health care in general, and home health care in particular. The percentage of forprofit home health care organizations increased from 6.7% of all HHAs in 1978 to 26.4% in 1983 (Estes et al. 1988). In his 1984 dissertation, Schlesinger (as reported in Marmor et al. 1987, p. 227) shows that for the three to five years before Medicare began covering home health services in 1981, proprietary agencies had 7% of the market share; for the three to five years after Medicare extended coverage that

percentage rose to 25%. Kane (1989, p. 25) reports that from 1982 to 1985 forprofit market share in home health went from 17% to 30% of Medicare-certified agencies. And, as noted, the number of Medicare-certified, forprofit HHAs has increased.

Of course, corporate provision of health care began *before* prospective payment (see, e.g., Schlesinger et al. 1987, p. 27); it is also true, as Light (1986) says, that pecuniary motives are not new to U.S. health care. But we do not argue forprofits are entirely new to home health, only that their prevalence has increased and that this increase has brought new pressures to bear on HHAs (Schlesinger 1985). However, holding competition constant in our analyses did not help distinguish between organizations of different tax status. This suggests social, as opposed to purely economic, consequences of the greater prevalence of forprofit organizations. The scope of those consequences awaits more research.

At the same time, it is also probable that, as forprofits entered home health care, they entered organizational fields already imbued with expectations for appropriate behavior. In this way nonprofits helped set standards for strategies and structures, prompting forprofit conformity with those standards. In one of the few published works on home health, Liszewski and Griffith (1988) empirically support an argument that coercive isomorphism, as theorized by DiMaggio and Powell (1983; see also Brown 1985; Paradis and Cummings 1986) creates similarity among HHAs. However, our findings regarding generalists and nonchains suggest such arguments are cast too broadly. While prospective payment and increased prevalence of forprofits have undoubtedly impelled some organizations toward similarity, our evidence indicates those forces have not impinged on all home health organizations in the same way. Chains and specialists, in particular, are more similar than are nonchains and generalists. Perhaps chains and specialists respond to different environments than their counterparts.

In their review, DiMaggio and Anheier (1990, p. 150) conclude that, "given available evidence, one can conclude only that legal form *does* make a difference, but the difference it makes depends on the institutional and ecological structures of the industry in question." Our findings support their judgment, but we would make the statement stronger still—it is the institutional and ideological features of organizational fields that give any meaning to the category "tax status" and thence to "forprofit" and "nonprofit."

APPENDIX

TABLE A1

MEANS, STANDARD DEVIATIONS, AND SAMPLE SIZES FOR ALL VARIABLES

	Mean	SD	N*
Independent and control variables:			
Tax status	566	.497	173
Chain	347	477	173
Size	92	109	163
Herfindahl index	.015	008	173
Dependent variables.			
Medicare certification	913	.282	173
Service provision:			
No of services provided	5 25	1 20	173
Provides infusion therapy	884	321	173
Provides physical therapy	931	255	173
Provides aide services	.983	.131	173
Provides hospice services	468	500	173
Provides homemaker	497	501	173
Provides ventilator	.618	487	170
Provides social work	879	.328	173
Maintains waiting list	.069	.255	173
Client mix:			
% 65 and older	75 9	21 2	169
% low income	33 9	28.1	130
% minority	27.2	22 9	139
No of clients served each month	213	305	165
Contracting			
Contract with acute hospitals	.634	483	172
Contract with psychiatric hospitals	145	.353	173
Contract with nursing homes	.439	498	173
Contract with board and care	335	.473	173
Contract with other home health organizations	.526	501	173
Contract with senior centers	279	450	172
Contract with state health department	202	403	173
Contract with hospices	.439	498	173
Contract with health maintenance organizations	576	496	172
Contract with preferred provider organizations	.355	.480	172
Staffing.			
No. of home health aides	26.1	48 0	166
No. of registered nurses	34.8	54.4	168
No of LVNs or LPNs	28.0	81 1	98
No. of physical therapists	6 53	6 56	152
Revenue sources·			
% from Medicare	58.1	37 8	168
% from Medicaid	3 83	8 05	167
% from insurance	16.6	20 7	163
% from copayments	16.0	28 5	162
% from philanthropy	1.61	6.14	165
Government dependence	61 9	38 6	167

* All *N*'s are not the same because of missing values

REFERENCES

Alexander, Jeffrey S., Laura L. Morlock, and Blair D Gifford. 1988. "The Effects of Corporate Restructuring on Hospital Policy Making." *Health Services Research* 23 (2): 311–37

Bergthold, Linda A., Carroll L. Estes, Pamela Hanes Spohn, and James H. Swan. 1988 "Running as Fast as They Can Organizational Changes in Home Health Care." Paper presented at the American Sociological Association meeting, Atlanta.

Bergthold, Linda A., Carroll L. Estes, and Augusta Villanueva. 1990. "Public Light and Private Dark. The Privatization of Health and Social Services for the Elderly." *Home Health Care Services Quarterly* 11 (3/4): 7–33.

Brown, Lawrence D. 1985. "Technocratic Corporatism and Administrative Reform in Medicare." *Journal of Health Policy, Politics, and Law* 10, no 3 (Fall): 579–99.

Carroll, Glenn R. 1984. "Organizational Ecology." *Annual Review of Sociology* 10.71–93

DesHarnais, Susan I., et al. 1987. "The Early Effect of Prospective Payment System on Inpatient Utilization and the Quality of Care." *Inquiry* 24:716–22.

DiMaggio, Paul J., and Helmut K. Anheier. 1990 "The Sociology of Nonprofit Organizations and Sectors." *Annual Review of Sociology* 16.137–59.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48.147–60.

Ermann, Dan, and Jon Gabel. 1984 "Multihospital Systems· Issues and Empirical Findings." *Health Affairs* 3 (1)· 50–64.

Estes, Carroll L., and Robert R Alford. 1990. "Systemic Crisis and the Nonprofit Sector Toward a Political Economy of the Nonprofit Health and Social Services Sector." *Theory and Society* 19 (2): 173–98.

Estes, Carroll L., Juanita B. Wood, and associates. 1988 "Organizational and Community Responses to Medicare Policy: Consequences for Health and Social Services for the Elderly." Institute for Health and Aging, University of California, San Francisco.

Feder, Judith J., Jack Hadley, and S Zuckerman 1987. "How Did Medicare's PPS Affect Hospitals?" *New England Journal of Medicine* 317 (14): 867–72.

Freeman, John, and Michael T. Hannan. 1983. "Niche Width and the Dynamics of Organizational Populations." *American Journal of Sociology* 88:1116–45.

Friedman, Bernard, and Stephen Shortell. 1988. "The Financial Performance of Selected Investor-owned and Not-for-Profit System Hospitals before and after Prospective Payment " *Health Services Research* 23 (2): 237–67.

Gay, E. Greer, Jennie J. Kronenfeld, Samuel L Baker, and Roger L. Amidon. 1989 "An Appraisal of Organizational Responses to Fiscally Constraining Regulation: The Case of Hospitals and DRGs." *Journal of Health and Social Behavior* 30.41–55.

Glazer, Nona Y. 1988. "At Work, at Home. Technology, Corporate Interests and the Labor Process in the Expansion of Home Care in the U.S." Paper presented at the American Sociological Association meeting, Atlanta.

Gornick, Marian, and Margaret Jean Hall. 1988 "Trends in Medicare Use of Posthospital Care." *Health Care Financing Review,* annual suppl. (December), pp. 27–28.

Guterman, Stuart, and Allen Dobson. 1986 "Impact of the Medicare Prospective Payment System for Hospitals " *Health Care Financing Review* 7 (3). 97–114.

Guterman, Stuart, et al. 1988. "The First Three Years of Medicare Prospective Payment· An Overview." *Health Care Financing Review* 9 (3). 25–55

Hall, Marnie, and Judith Sangl. 1987. "The Impact of Medicare's PPS on Long-Term

Care Providers." Paper presented at the American Public Health Association meeting, New Orleans.

Hannan, Michael T., and John Freeman. 1977. "The Population Ecology of Organizations." *American Journal of Sociology* 82 (5): 929–64.

———. 1984 "Structural Inertia and Organizational Change." *American Sociological Review* 49:149–64

———. 1989. *Organizational Ecology*. Cambridge, Mass.: Harvard University Press

Hansmann, Henry. 1987. "Economic Theories of Nonprofit Organization." Pp. 27–42 in *The Nonprofit Sector: A Research Handbook,* edited by Walter W. Powell. New Haven, Conn.. Yale University Press.

Harrington, Charlene. 1984. "Public Policy and the Nursing Home Industry." Pp. 144–54 in *Readings in the Political Economy of Aging,* edited by Meredith Minkler and Carroll L. Estes. Farmingdale, N.Y.: Baywood.

Hirschman, Albert O. 1964. "The Paternity of an Index." *American Economic Review* 54:761–62.

Hollingsworth, J. Rogers, and Ellen Jane Hollingsworth. 1989. "Public Organizations: The Behavior of Hospitals in England and Wales." Pp. 110–36 in *The Nonprofit Sector in International Perspective: Studies in Comparative Culture and Policy,* edited by Estelle James. New York. Oxford University Press.

James, Estelle 1987 "The Nonprofit Sector in Comparative Perspective." Pp 397–415 in *The Nonprofit Sector: A Research Handbook,* edited by Walter W. Powell. New Haven, Conn.. Yale University Press.

———. 1989. Introduction. Pp. 3–27 in *The Nonprofit Sector in International Perspective: Studies in Comparative Culture and Policy,* edited by Estelle James. New York. Oxford University Press.

Kane, Nancy M. 1989. "The Home Care Crisis of the Nineties." *Gerontologist* 29 (1): 24–31.

Leader, Shelah. 1987. "Meeting the Home Health Care Needs of the American Elderly: Policy Dilemmas and Muddling Through " Paper presented at the American Public Health Association meeting, New Orleans.

Levitz, Gary S., and Paul P. Brooke, Jr. 1985. "Independent versus System-Affiliated Hospitals: A Comparative Analysis of Financial Performance, Cost, and Productivity." *Health Services Research* 20 (3): 315–37.

Light, Donald W. 1986. "Corporate Medicine for Profit." *Scientific American* 255 (6): 38–45.

———. 1991. "The Restructuring of the American Health Care System." Pp 53–65 in *Health Politics and Policy,* 2d ed. Edited by Theodor Litman and Leonard S. Robins. New York: Delmar.

Liszewski, Diane M., and Elsie I. Griffith. 1988. "Diversification and Corporate Restructuring Revisited." *Nursing Clinics of North America* 23 (2): 399–413.

Luke, Roice D., James W. Begun, Dennis D. Pointer. 1989. "Quasi Firms: Strategic Interorganizational Forms in the Health Care Industry." *Academy of Management Review* 14 (1): 9–19.

Marmor, Theodore R., Mark Schlesinger, and Richard W. Smithey. 1987. "Nonprofit Organizations and Health Care." Pp. 221–39 in *The Nonprofit Sector,* edited by Walter W. Powell. New Haven, Conn.. Yale University Press.

Meyer, Marshall W. 1987. "The Growth of Public and Private Bureaucracies." *Theory and Society* 16:215–35.

Meyer, Marshall W., William Stevenson, and Stephen Webster. 1985. *Limits to Bureaucratic Growth*. New York: Walter de Gruyter.

Neu, C. R., and Scott C. Harrison. 1988 "Posthospital Care before and after the Medicare PPS." Report to Health Care Financing Administration, no. R-3590-HCFA. Santa Monica, Calif.: RAND.

Newcomer, Robert, Juanita Wood, and Andrea Sankar. 1985. "Medicare Prospective Payment: Anticipated Effect on Hospitals, Other Community Agencies, and Families." *Journal of Health Politics, Policy and Law* 10 (2): 275–82.

Olson, William B., and Karen K. Olson. 1988. "Interorganizational Relationships of Hospitals and Home Health Agencies in Arkansas " Unpublished manuscript. Arkansas State University, College of Business.

Paradis, Lenora Finn, and Scott B. Cummings. 1986. "The Evolution of Hospice in America toward Organizational Homogeneity " *Journal of Health and Social Behavior* 27:370–86.

Pattison, Robert V., and Hallie M. Katz 1983. "Investor-Owned and Not-for-Profit Hospitals." *New England Journal of Medicine* 309 (6). 347–53.

Perry, James L , and Hal G. Rainey. 1988 "The Public-Private Distinction in Organization Theory: A Critique and Research Strategy." *Academy of Management Review* 13 (2): 182–201.

Powell, Walter W , and Rebecca Friedkin. 1987 "Organizational Change in Nonprofit Organizations " Pp. 180–94 in *The Nonprofit Sector: A Research Handbook,* edited by Walter W. Powell. New Haven, Conn.: Yale University Press.

Prospective Payment Assessment Commission (PROPAC) 1990 "Medicare Prospective Payment and the American Health Care System " Report to Congress, no. 621 Chicago: Commerce Clearing House, Inc.

Rainey, Hal G., Robert W. Backoff, Charles H. Levine. 1976. "Comparing Public and Private Organizations " *Public Administration Review* 36, no 2: 233–44

Relman, Arnold S. 1983. "Investor Owned Hospitals and Health Care Costs." *New England Journal of Medicine* 309 (6)· 370–72.

Ritzer, George, and David Walczak. 1988. "Rationalization and the Deprofessionalization of Physicians." *Social Forces* 67 (1): 1–22.

Salamon, Lester M. 1987. "Partners in Public Service. The Scope and Theory of Government-Nonprofit Relations." Pp. 99–117 in *The Nonprofit Sector: A Research Handbook,* edited by Walter W. Powell New Haven, Conn.: Yale University Press.

Salvatore, Tony. 1981. "Interorganizational Survival Strategies: Options for the Voluntary Nonprofit Home Health Agency." *Home Health Review* 4 (2). 14–18

——— 1982 "Diversification: Viable Strategy for the Voluntary Nonprofit Home Health Agency?" *Home Health Review* 6 (4): 31–37.

Scherer, Frederic M. 1970 *Industrial Market Structure and Economic Performance.* Chicago. Rand McNally

Schlesinger, Mark. 1985. "The Rise of Proprietary Health Care." *Business Health* 2 (1): 7–12

Schlesinger, Mark, and R. Dorwart. 1984. "Ownership and Mental Health Services A Reappraisal of the Shift toward Privately Owned Facilities " *New England Journal of Medicine* 311:959–65.

Schlesinger, Mark, et al. 1987 "The Privatization of Health Care and Physician's Perceptions of Access to Hospital Services." *Milbank Quarterly* 65 (1) 25–58.

Schramm, Carl J., and Jon Gabel. 1988. "Prospective Payment Some Retrospective Observations." *New England Journal of Medicine* 318:1681–83.

Shamansky, Sherry L. 1988. "Providing Home Care Services in a For-Profit Environment." *Nursing Clinics of North America* 23 (2). 387–98.

Shortell, Stephen M. 1988. "The Evolution of Hospital Systems: Unfulfilled Promises and Self-Fulfilling Promises." Arthur Weissman Memorial Lecture, University of California, Berkeley, School of Public Health, April 20.

Shortell, Stephen M., et al. 1986. "The Effects of Hospital Ownership on Nontraditional Services." *Health Affairs* 5 (Winter): 97–111.

Spohn, Pamela Hanes, Linda Bergthold, and Carroll L. Estes. 1988. "From Cottages

to Condos: The Expansion of the Home Health Care Industry under Medicare." *Home Health Services Quarterly* 8 (4): 25–55.

Starr, Paul 1982. *The Social Transformation of American Medicine*. New York Basic.

Steinberg, Richard 1987. "Nonprofit Organizations and the Market." Pp. 118–40 in *The Nonprofit Sector: A Research Handbook,* edited by Walter W. Powell. New Haven, Conn.: Yale University Press.

Swan, James H., and Carroll L. Estes. 1990. "Changes in Aged Populations Served by Home Health Agencies." *Journal of Aging and Health* 2 (3): 373–94.

Swan, James H., Adela de la Torre, and Rachel Steinhart. 1990. "Ripple Effects of PPS on Nursing Homes: Swimming or Drowning in the Funding Stream?" *Gerontologist* 30 (3). 323–31.

Szasz, Andrew. 1990 "The Labor Impacts of Policy Change in Health Care—How Federal Policy Transformed Home Health Organizations and Their Labor Policies " *Health Policy, Politics, and Law* 15 (1): 191–210.

Trager, Brahna. 1990. "Home Care and Public Policy." *Caring* 9 (March)· 4–7.

U.S. Congress. 1983. *Social Security Amendments of 1983 Prospective Payments for Medicare Inpatient Hospital Services, Title VI, PL 98-21* Washington, D.C : Government Printing Office.

U.S. House of Representatives, Committee on Ways and Means. 1990. *Overview of Entitlement Programs: 1990 Green Book*. 101st Cong., 3d sess. Committee Print.

Vladeck, Bruce C. 1980. *Unloving Care: The Nursing Home Tragedy*. New York. Basic.

Waldo, Daniel R., Katharine R. Levit, and Helen Lazenby. 1986. "National Health Expenditures, 1985." *Health Care Financing Review* 8 (1). 1–21.

Weisbrod, Burton A. 1988 *The Nonprofit Economy*. Cambridge, Mass.. Harvard University Press.

——— 1989. "Rewarding Performance That Is Hard to Measure. The Private Nonprofit Sector." *Science* 244 (May 5). 541–46.

Wood, Juanita B., and Carroll L. Estes 1990. "The Impact of DRGs on Community-based Service Providers. Implications for the Elderly." *American Journal of Public Health* 80 (7)· 840–43

Zucker, Lynne G. 1988. *Institutional Patterns and Organizations: Culture and Environment,* edited by Lynne G. Zucker. Cambridge: Ballinger

Restructuring Patterns of Elite Dominance and the Formation of State Policy in Health Care[1]

Allen W. Imershein, Philip C. Rond III, and Mary P. Mathis
Florida State University

The authors analyze recent state intervention to control health care costs as both reflecting and fostering changes in previously prevailing patterns of dominance among health care elites and between those elites and nonhealth elites. The authors contend that these health policy outcomes reflect a divergence of material interests and a political fragmentation among dominant health elites and between those and nonhealth, especially corporate, elites, which previously allowed health provider elites to control the health policy agenda. Empirical analysis shows that patterns of fragmentation among hospital elites and cohesion among business elites, in the form of business coalition development, can account for the adoption/retention of the four predominant state health care cost containment policies during the period 1981–87. The authors conclude that a political and economic restructuring of elite dominance is taking place but it is occurring unevenly across states and regions.

The substance and impact of recent revisions in federal Medicare policy and numerous changes in state health policies have drawn the attention of the health care industry, policymakers, and scholars. Much has apparently changed since Alford (1975) scrutinized the "symbolic" health care politics more than 15 years ago. While the debate over federal changes has been fairly concentrated and heated, the discussion about state policies has been much more diffuse, owing in part to the diversity and unevenness of the policy changes made at the state level. The research discussed herein contends that examining the diversity and unevenness

of state-level health policy changes may prove more indicative of the extent to which, and the ways in which, fundamental and potentially long-term restructuring of health care market relations—and, especially, the arrangements of power and influence within the American health care arena—may be taking place. To consider these issues we examine health cost containment policies passed by the states, from 1981 through 1987, in relation to the varying sets of elite relations that structure health care markets.

Much of the recent literature, which grows out of the network analytic tradition of power-structure research, has dealt with the substantive issue of establishing that sufficient cohesion exists within a group or class for it to exercise its collective will (see Useem 1983, 1984; Laumann and Knoke 1987). The present research follows in this tradition of power structure studies. However, instead of searching for stable power arrangements within a network of social relations, we attempt to identify disruptions that lead to a fragmentation within established networks of elite relations. This structural change occurs via a process triggered by the divergence of the material interests of key elites.

We propose that when the presumptive benefits of the prevailing arrangement are not sufficiently realized, elites will actively seek to modify structural arrangements (i.e., realign structural power) better to serve their particular interests. As underlying traditional economic and political relations weaken or break up (i.e., fragment), new coalitions will emerge and press for an expansion of the policy agenda to accommodate their interests (cf. Fine 1984). Policy outputs will in turn reflect this divergence and foster new structural accommodations. These new accommodations will be encapsulated in alterations in the state administrative/bureaucratic apparatus designed to legitimate and perpetuate the new set of accommodations. As employed throughout this research, the term "elite fragmentation" is intended as a shorthand descriptor of the change process sparked by the breakdown of structural accommodations to the material interests of key elites.

Up until the past two decades studies of elite fragmentation within the health care arena were largely moot. From the early 1930s—when the medical profession established its hegemony over health affairs—until the early 1970s, health sector arrangements were governed by a broad, implicit alliance among both health and nonhealth elites, premised chiefly upon accommodation to the interests of the medical profession (Starr 1982), and secondarily to the interests of the voluntary hospitals. Developments following in the wake of the 1960s federal health policies intended to further accommodate dominant structural interests, instead undermined them.

Specifically, we will argue that economic and political developments

occurring between 1965 and 1980 produced numerous cleavages within and between concerned elites and inexorably led to a combined fiscal and legitimation crisis within American health care. This crisis in turn spawned the formation of new elite alliances, the expansion of state and federal health policy agendas during the 1980s, and the adoption of national and state health policies aimed at altering elite relations in and around health care.

Our arguments depart from previous theories and arguments that focus only at the level of policy-making—essentially pluralist arguments that attribute too great a significance to symbolic politics and assume a direct relation between that activity and the underlying arrangements of power (see Alford and Friedland 1985). We also go beyond arguments that consider the economic impetus behind policy-making, but that focus predominantly (and almost exclusively) on the entry of nonhealth corporate elites to account for recent state policy changes (see esp. Bergthold 1990). The latter arguments are more congruent with the ones presented here in recognizing the reconfiguration of the political arena. However, they are limited by insufficient examination of that reconfiguration and its relation to the market arena, a matter central to assessing the degree of rearrangements of power. Our approach attempts to account more adequately for both the complexity and variability of the emerging patterns of health care policy and power. To address these issues we pose several theoretical and empirical arguments.

Argument 1

A significant, albeit varying, divergence of economic/material interests has developed among health elites and between those elites and other nonhealth elites.

The implicit alliance of health and nonhealth elites, which has dominated the American health care arena (Freidson 1970; Stevens 1971; Starr 1982), has often been referred to as a medical-industrial complex (Ehrenreich and Ehrenreich 1970; Relman 1980). That alliance included (1) the health care provider elites, including (a) the medical profession, which demonstrated its political clout through the American Medical Association, and (b) the hospitals, which acted on their own behalf through the American Hospital Association; (2) insurers, particularly Blue Cross/Blue Shield—formerly considered health elites but now most often acting as external elites/purchasers of care; (3) fractions of industrial capital (nonhealth elites), which exerted an economic influence through gifts and endowments predominately to the voluntary hospital sector and the teaching hospitals of the industrial Northeast and Midwest; and (4) spe-

cific sectors of capital that produced health care technology, such as pharmaceuticals, medical equipment, and hospital supplies.

The patterns of dominance of the health care system were constituted in large part through the workings of market relations designed to serve the material interests of the participants, especially those of health care providers. "Serving the material interests" here entails not only maintaining the production of visible economic benefits and advantages for certain elites, but also maintaining the production/reproduction of the market and the broader structural conditions under which elites have sustained their position of economic advantage: that is, controlling conditions and definitions of both work and trade, limiting access to markets, limiting or controlling relevant statutory regulations, controlling training and licensing, and so on. Thus, physicians initially conducted all health care activity and, through controlling entry into their own ranks, controlled overall market entry. The development of voluntary hospitals as the locus of more intense and/or technologically based health care was complemented by Blue Cross/Blue Shield in its role as financial underwriter of labor and capital costs. Likewise, Medicare's reimbursement practices, adopted from "the Blues," served the economic interests of the physicians and the hospitals.

In general, through the 1970s, nonhealth elites, enjoying the benefits of economic expansion, granted health provider elites political and organizational control of the health care system. (For one argument on the capital-medical profession relationship, see Brown [1979]; for a detailed examination of this history, see Light [1991]). So long as the economy expanded and the set of elite relationships remained relatively stable and closed, the material interests of the respective elites were, for the most part, being met, the health care market was rarely recognized or designated as such, and both the material interests of the respective elites and their respective political expressions were commonly shared.

Recent economic and political events have precipitated a divergence of interests among formerly allied elites. These include the massive inflation of the late 1970s coupled with declining corporate profits, recession, and continuing health care inflation; the growth of forprofit hospital chains; changes in federal Medicare reimbursement and other related tax policies; and prevailing concerns for cost control. Three of these are particularly noteworthy.

First, the health care implications of the major regional and sectoral reconstruction of the American economy, which began during the 1970s (Bluestone and Harrison 1985; Davis 1984), were most strongly felt in the urban Northeast and the Midwest, where the voluntary hospitals had dominated health care since the 1930s. The economics of the 1970s

coincided with federal policies to encourage the growth of the forprofit hospital sector and the emergence of the forprofit hospital chains. This occurred primarily in the Sunbelt, where population growth was greatest and voluntary hospitals comparatively less developed (Markowitz and Rosner 1973; Starr 1982). The result was a significantly regionalized hospital system and a potentially restructured health care system.

Second, during the 1970s corporate elites became increasingly alarmed over health care costs that were rising at a rate well in excess of the general inflation rate. International competition, coupled with the restructuring of the American economy, had produced declining profits for many of these corporate purchasers of health care, exacerbating this cost rise. Through the late 1970s and early 1980s, corporate elites, led by the Washington Business Group on Health, set about creating the institutional mechanisms (i.e., autonomous business coalitions for health action) necessary to challenge health provider dominance of the health policy agenda at both the federal and state levels. By the time the nation slipped into the recession of the early 1980s, corporate elites were prepared to engage in the business of health care policy-making.

In addition, federal policy initiatives during the 1970s and 1980s helped to shift the asymmetrical relation between the nonprofit hospital sector and the forprofit sector in favor of the latter (Hoff and Schaner 1982). Medicare reimbursement policy (Starr 1982) made the industry attractive to investors. The 1974 National Health Planning and Resources Development Act, intended to regulate the supply of hospital beds by requiring a certificate of need (CON) review, sanctioned the construction of vertically integrated systems, thus stimulating the formation of multi-institutional hospital chains (Salkever and Bice 1976: Bergstrand 1982; McKinlay and Stoeckle 1988).[2] The tax policies of the procompetition Reagan administration eroded the advantage of tax-exempt status for the nonprofits (Hoff and Schaner 1982). The overall effect of these policies was to stimulate a fragmenting of the hospital industry.

Argument 2

The divergence of material interests has resulted in a fragmentation of previous implicit and explicit institutional, organizational, and political alliances and has resulted in the loss of unified control of the health policy agenda by health provider elites.

---

[2] In limiting new hospital construction, the act increased the value of existing facilities, thereby making these facilities financially attractive to expanding investor-owned enterprises. Further, it not only placed no restrictions on vertical multihospital expansion but contained language specifically encouraging such growth.

The once highly controlled political arena has become one in which there is competition and often conflict among formerly allied health and nonhealth elites. Hence, recent politics and policies are different from those of the "dynamics without change" period of the 1970s, when apparent conflict over costs merely represented efforts to maintain the political and professional ideology that legitimated health-care-provider elite dominance (see Alford 1975; Imershein and Rond 1989). Because health care policy-making during the 1970s continued to be dominated by health provider elites, early efforts to contain health care costs merely maintained the provider-biased cost-based reimbursement system. The failure of incremental efforts to contain costs coupled with the growing economic crisis began to undermine the legitimacy of the prevailing structural arrangements (see Imershein [1977] for an earlier version of this argument) and contributed to a growing fragmentation among previously allied health and nonhealth elites.[3] Under these conditions of fragmentation during the 1980s, one or more elites have sought to have their interests served through state intervention via health care policy-making.[4] The resulting political landscape, we argue, may be systematically characterized in terms of patterns of fragmentation and/or cohesion among and across these elites. That is, the degree of influence on, or control over, the policy agenda is dependent upon the degree to which the respective elites are fragmented or cohesive.

Argument 3

The specific health care policy, or policies, chosen by a state during the period 1981–87, reflects the particular patterns of health elite dominance in the state, the ways in which the material interests of these dominant elites may have diverged, the emergent conflicts between health elite interests and those of nonhealth elites seeking control in/over the health care sector, and which of these elites have been able to control the health policy agenda. These policies in turn foster a set of market relations that best serves the interests of that elite group, or coalition across elite groups, that establishes or maintains control over the health policy agenda in the state.

[3] This fracturing of previously aligned elite interests (i.e., elite fragmentation) may, for example, involve: (1) dissension among the producers of goods and services within the sector, (2) cleavages between producers and purchasers, and/or (3) divisions between different purchaser groups (Burt 1983). One or another elite may seek state intervention on its own behalf in the name of broader societal goals.

[4] For more on the possible results of how unresolved fragmentation of interests may create opportunities to expand the public policy agenda to include previously excluded problems and solutions, see Imershein and Rond (1989).

In the wake of the "New Federalism" initiated by the Reagan administration, state legislatures increasingly addressed issues of health care cost containment (see Lee and Estes 1983). The 1981 Omnibus Budget Reconciliation Act officially devolved greater authority over health programs to the state level while cutting state funding levels. Although the actual extent of authority shifts from federal to state proved to be less extensive than was first claimed, the rhetoric accompanying and justifying the shift clearly legitimated greater assertions of authority at the state level (Ellwood 1982; Imershein and deHaven-Smith 1987). The concurrent shift in health sector economic battles, reflected in the development of state-level business health coalitions and forprofit hospital trade associations, combined with this greater authority to result in a set of varying health cost containment policies.

GENERAL EXPECTATIONS ABOUT ELITE RELATIONS AND MARKET ARRANGEMENTS

From this point in our discussion we move forward at two levels of analysis. We elaborate our theoretical formulations that relate health and nonhealth elites. And we expirically examine what we believe to be the most important (but by no means sole) concrete presentation of health elite fragmentation (i.e., within the hospital industry), the most significant evidence of a cohesive nonhealth elite (i.e., business health coalitions), the way in which health and nonhealth elites relate (i.e., in the health care market), and the most revealing evidence of changing relations among elites (i.e., state health care policies) during the 1980s. We make no empirically grounded claims beyond this specific instance of health elite fragmentation/nonhealth elite cohesion. At the same time, we believe the more abstract theoretical formulation has broader import than the single, albeit central, example to which we direct our attention here, and therefore we present both abstract and concrete formulations where possible.

We see four patterns of elite fragmentation/cohesion that will in turn produce, under the empirical conditions we theoretically propose, four distinct market models (see fig. 1). In the specific case of hospital elites and business coalitions the respective patterns of fragmentation/cohesion will result in one of four policy alternatives that will in turn foster (or be intended to foster) corresponding market relations (see fig. 2).

1. *Unified health elite/fragmented nonhealth elite.*—Where health elites are able to resist fragmentation and have little challenge from cohesive external elites, market relations are expected to continue to take the form of supply regulation, maintaining the advantage of existing providers/suppliers, and health care policies will be chosen to foster these ar-

Fragmentation Within Health Sector

		LOW	HIGH
Cohesion Outside Health Sector	LOW	supply regulation	open competition
	HIGH	price and product regulation	negotiated competition

FIG. 1.—Patterns of health sector fragmentation, degree of external nonhealth cohesion, and expected market relations

Fragmentation Within Hospital Industry

		LOW	HIGH
Development of Business Coalitions	LOW	maintain CON	repeal CON
	HIGH	rate setting/ review	hospital financial disclosure

FIG. 2.—Patterns of hospital industry fragmentation, development of business coalition activity, and expected health care policies.

rangements. Empirically, under those conditions by which the hospital industry is still controlled by the voluntary hospital sector, that sector would maintain its market control and limit market entry by seeking to retain CON policies intended to restrict market entry and to block efforts to impose external controls on hospital revenues.

2. *Unified health elite/cohesive nonhealth elite.*—Where health elites are able to resist fragmentation but are challenged by cohesive external elites, market relations are expected to take the form of price and product regulation, maintaining the market position of present providers but subjecting them to government-administered price controls in response to external demands. Empirically, where a unified hospital industry (again under the dominance of the voluntary sector) is challenged by powerful business interests seeking to control costs, health care policies will be chosen that constrain hospital pricing practices but still maintain market entry control; that is, rate setting or rate review.

One might argue that low fragmentation of hospital elites combined with high cohesion of business elites would result in a situation whereby negotiated competition would result, that is, big buyers and big sellers. While possible, it is likely only where business elites wield more power than is ordinarily the case with a cohesive hospital industry's still having control of the health policy agenda; this is unlike the case with a fragmented hospital industry considered below, which does, we argue, produce negotiated competition. Although it somewhat limits market rewards, price and product regulation allows greater market predictability and is generally more favorable toward continued market control by the hospital industry than negotiated competition is likely to be (alternative 4 below). And, in the case of either price and product regulation or negotiated competition, while cohesive business elites may be sufficiently powerful to gain state intervention in the form of new policies to improve their market position, they have never been sufficiently unified or powerful in the market itself (at least so far) to dictate terms to hospital elites.

3. *Health elite fragmentation/nonhealth elite fragmentation.*—Under conditions of fragmentation among health provider elites, the challenging provider elite will seek to enhance its material interests; that is, it will seek policy support to increase opportunities for market entry or expansion, in other words, the repeal of CON. This will occur where conditions are marked by the significant entrance of investor-owned hospitals into the health care market, but where business coalitions seeking to control costs are absent or underdeveloped; market relations will then take the form of open competition.

4. *Health elite fragmentation/nonhealth elite cohesion.*—Where health elites have fragmented and at the same time are faced with challenges from external elites, market relations are expected to take the form of negotiated competition (of prices/rates). These relations would favor the market position of large purchasers of care (corporate business entities as external elites), and health care policies would then be chosen to foster these arrangements through means such as increasing the availability of comparative hospital pricing/hospital costs, fostering HMO and/or PPO development, and, especially, promoting the passage of hospital financial disclosure policies.

Under conditions of economic strain, such as declining revenues and increasing health care costs, the state may find itself in a position similar to that of other large purchasers of care and reconsider its long-standing tendency to concur with provider dominance of health policy. Accordingly, agents of the state may act in concert with or in a manner similar to those purchaser elites and favor cost containment policies, aimed at changing health market relations, that enhance the state's power to regulate or negotiate control over health sector revenues.

MEASUREMENT OF ELITE FRAGMENTATION/COHESION AND
STATE POLICY ALTERNATIVES

Health/Hospital Elite Fragmentation

The fragmentation occurring among and between health elites takes
many forms. Divisions have occurred among three notable health elite
groups: physician elites, hospital elites, and Blue Cross and other health
insurance elites. Most notable is the division *within* the hospital industry
itself. As others have noted, the common position of forprofit and non-
profit hospitals within the delivery of health care necessarily entails some
overlapping material interests, and hospitals' actual day-to-day perfor-
mance may appear in many ways more similar than different. Nonethe-
less, the development of significant numbers of forprofit hospitals in a
given state has resulted in both economic and political divisions within
the state's hospital industry as a whole (see Imershein and Rond 1989).

Data published by the American Hospital Association in 1972 are con-
sistent with our assumption of minimal fragmentation: on average, only
6.5% of all community hospital beds in the United States were reported
to be under proprietary ownership at that time. By the end of 1986 the
number of forprofit beds in the nation had increased by 68.8%. For
example, the percentage of forprofits in Nevada grew by 60.6%; in Texas
the rate of increase was 62.9%; in Georgia the percentage increased by
134.0%; and in Tennessee by 126.7%. The largest increase was in Flor-
ida, where the presence of forprofit hospitals grew by 288.0%.

As evidence of these divisions and reflecting the growth and differentia-
tion of the hospital industry, new hospital trade associations emerged to
represent the particular interests of various segments of the industry: the
National Council of Community Hospitals, the Voluntary Hospitals of
America, and, representing forprofit hospitals, the Federation of Ameri-
can Hospitals (*Encyclopedia of Associations* 1988). The emergence of
these organizations not only reflects the growing divergence of economic
interests within the industry, but also marks the decline of hospital indus-
try solidarity in the national political arena.

To demonstrate health elite fragmentation only in terms of the hospital
industry is not, we believe, to suggest that no other economic divergences
among health elites has occurred. For example, Blue Cross often now
acts primarily in its role as a large purchaser of care, and, as such, the
material interests of what had been one of the traditional health elites
now often diverge significantly from those of hospitals as providers. Un-
der such conditions, Blue Cross more likely works in concert with other
large purchasers of care (and, therefore, against hospitals). Moreover, the
presence of significant fragmentation within the health insurance industry
itself is also likely. Unfortunately, proprietary control of access to data

on changing market shares within the health insurance industry precludes their use in our analysis. Further, while fragmentation could also be considered along a number of other lines among health elites in general or hospital elites in particular, the health policy alternatives we consider appear to have the most widespread importance for changing market relations among hospital elites, and the breakout of the forprofits within the hospital industry to have a very significant if not consistently the greatest importance for overall health care markets. We focus, therefore, on data from the hospital industry for our empirical analysis of health elite fragmentation.

We examined hospital data at three different points in time (1972, 1980, 1986) and considered the implications of various measurement alternatives for assessing the propensity to adopt alternate policy and market arrangements. We concluded that, in some instances, the rate of growth in the number of beds in proprietary hospitals might be more relevant, while in others the absolute proportion of forprofit beds might be more decisive in prompting divisiveness of economic and political interests within the industry. In either event, we believe that for the purposes of demonstrating the general applicability of our theoretical approach it is appropriate to characterize states as having high or low fragmentation within the hospital industry by using a split mean approach based on the percentage of forprofit (investor-owned) community hospital beds, by state, in 1980.[5] This is being examined after the formation of most investor-owned hospital chains and thus after the initial fragmentation in the hospital industry; it is also after the adoption of federally mandated CON regulations, but before the adoption of most new state-level health cost containment policies.

Cohesion of Nonhealth Elites

In contrast to analyses such as those of Laumann and Knoke (1987), which seek to demonstrate the cohesion of elites (in our terms, "nonhealth" elites) by using network analysis, we assume a general unity of nonhealth (business) elites with regard to health care. The development of business health coalitions then signaled for us that the *implicit* sets of relations relative to health care among health and nonhealth elites, which had previously served the latter's material interest, no longer accom-

[5] As can be seen from table 1 below, all states with 7% or more forprofit beds were classified as high fragmentation states In examining alternative formulations of this variable (e g., percentage of total beds under forprofit ownership vs. percentage of growth of forprofit beds vs. various combinations of these two dimensions) and alternative classification rules (e.g., mean vs median split), we obtained similar analytic results.

plished that end. If corporate leaders (nonhealth elites) had previously been unconcerned with issues of health policy for economic reasons or because of some tacit agreement with health elites, the emergence of business coalitions during the early 1980s clearly demonstrated that elite relations between health and nonhealth elites *and* among nonhealth elites themselves had changed. Now corporate leaders *openly* joined with each other and *collectively* challenged health elites in the health policy arena. Such activity, we claim, reveals a special case of nonhealth elite cohesion. In addition, in some states organized labor, the elderly (through the American Association of Retired Persons, or the AARP), or other "grass-roots" coalitions took the lead in espousing the adoption of health policies adverse to the interests of health elites and favorable to the interests of one influential group of nonhealth elites: corporate purchasers of health care.

We wish to avoid the problems attendant to measuring structural power via indicators of pluralist participation, so we have elected to measure nonhealth elite cohesion in terms of the extent to which business elites succeeded in establishing the "institutional mechanism" deemed necessary to challenge health provider control of state health policy agendas. As noted earlier, national business leaders believed that autonomous state and local business coalitions for health action were needed to mobilize the bias of decision making in a direction more favorable to the interests of corporate health care purchaser elites.

Using the 1984 and 1985 editions of the *Directory of Business Coalitions for Health Action,* we identified business coalitions in each of the 50 states and assigned the following ratings in order to assess their autonomy and the extent of their organizational development.[6]

Corporate status: unincorporated = 0, incorporated = 1;

Staffing: volunteer = 0, paid = 1;

Revenue: none or donations = 0, dues, grants, fees = 1;

Activities: education only = 0, other activities (excluding lobbying) = 1, other activities (including lobbying) = 2.

States were classified as having a high degree of external (nonhealth) elite cohesion if they had one or more autonomous coalitions that scored

[6] The years 1984 and 1985 were chosen as representing the period of peak coalition development. Other sources indicate that after 1985 very few new coalitions developed and that a number of existing coalitions either became dormant, ceased operations, or experienced a change of mission

at least four points on the organizational development scale. Those coalitions whose membership was reported to consist of less than 5% health care provider representatives (i.e., physicians and hospitals) were deemed to be autonomous from health elites. In all, 28 states received a classification of high cohesion; the remaining 22 received a classification of low cohesion. Using the respective categorizations of fragmentation and cohesion across the states, we were then able to examine the relation between the patterns of fragmentation/cohesion in each state and the health policies chosen by those states.

State Policy Alternatives

In order to assess the adequacy of our framework we examined the actions of each state relative to four cost containment policy alternatives that foster alternate sets of market relations.

1. *Retain certificate of need (only)*.—State regulation of the supply of hospital (and some other health care) facilities was implemented during the 1970s in the 49 states that complied with the 1974 National Health Planning and Resources Development Act by establishing CON programs, which were intended to contain health costs through supply regulation/control.[7] Retention of CON (or other closely related mechanisms),[8] absent the adoption of one of the other two focal policies (items 3 and 4 below), is viewed as evidencing a continued commitment to market relations based upon supply regulation. A state may adopt another policy while retaining (or repealing) CON, but we judge those alternatives as either more constraining (in the case of rate review or rate setting) or more competitive (in the case of financial disclosure) and thus classify states according to those policies as described below.

2. *Repeal the certificate of need program (only)*.—Policies encouraging open competition among hospitals would be sought in states where for-profit hospital elites had captured a significant market share but still felt so constrained by supply regulation that they were willing to disrupt the traditional political unity of the hospital elites and seek redress in the political arena. This activity most typically would consist of efforts to repeal CON. Thus, policies favoring competition are unlikely to be passed unless a significant degree of fragmentation within the hospital industry (and underlying divergence of interests) has *already* occurred, regardless of the rhetoric of competition that might be voiced outside the hospital industry.

3. *Hospital rate setting/rate review*.—Price and product regulation,

[7] Louisiana had a less stringent form of supply regulation.
[8] These would include hospital construction and/or licensure moratoria/quotas

that is, hospital rate setting or rate review, should be dominant in those situations where hospital elites share a common set of material interests, but where these interests are in opposition to those of cohesive business elites. This policy seeks to preserve the market entry control of dominant hospital elites while at the same time attempting to accommodate the cost control interests of cohesive business elites. Hospital rate setting dates from the late 1960s and early 1970s, when six states, five located in the Northeast, adopted mandatory hospital rate review in an attempt to contain rapidly rising hospital costs. More recently, in 1983, price and product regulation was implemented in Medicare's diagnostic related groups (DRG) program. In this situation, the federal government exercised its power as the nation's largest single purchaser of health services and imposed an administered price structure and uniform product definitions upon the hospital industry. In so doing, the federal government was responding to general concern and to calls for cost control chiefly from the business/corporate community. This federal initiative coincided with the emergence of politically active health-oriented business coalitions throughout the nation (Bergthold 1987). The federal government had also been under considerable fiscal strain as the result of Reagan administration tax cuts and the recent recession. It was thus acting significantly in the role of a cohesive external elite/purchaser of care in imposing price and product control.

   4. *Hospital financial disclosure (without rate setting).*—Recently, large employers and others have actively sought to negotiate price and product with hospital providers. By consolidating purchasing volume within the health care market, these elites possess the potential to demand preferred pricing and product arrangements. In order to achieve the conditions necessary to market arrangements based upon this form of elite negotiation, business elites must have access to hospital financial data. State health policy that supports such arrangements primarily includes legislation requiring some form of hospital financial disclosure.[9] Financial disclosure may also occur as part of a rate-setting policy, that is, as a prerequisite to being able to set rates; but we regard rate setting as constraining/regulatory rather than competitive and therefore classify states that require financial disclosure as a part of rate-setting only under the latter category.

   Negotiated competition differs from open market competition in that a smaller number of parties participate directly in setting the conditions for market transactions. It differs from regulatory approaches in that there are fewer formal barriers to market entry and no unilateral ability

[9] Other related policies may foster development of health maintenance organizations (HMOs) and preferred provider organizations (PPOs).

to set the conditions of price and product. Arrangements based principally upon negotiated competition are, however, often compatible with strong regulatory policies (e.g., Florida's rate review system).

We reviewed documents published by the National Conference of State Legislators (NCSL 1985; Polchow 1986; Pierce and Hekman 1984; King and Frances 1987), the Intergovernmental Health Policy Project (1985), the National Association of Health Data Organizations (NAHDO), and the American Health Planning Association, then supplemented this information with contact with staff members of these organizations, to identify states that, through the 1987 legislative session, had adopted, retained, or repealed each of the four cost-containment policies listed above. In the sections that follow, we examine the policy-adoption behaviors of the states within the context of each of the four market models.

RESULTS OF RESTRUCTURING· NEW MARKET RELATIONS AND STATE POLICY IN HEALTH CARE

Our theoretical framework, based upon the specific configuration of fragmentation among hospital elites and the development of business coalitions within each state, leads to several expectations for the 43 relevant states during the period under study:[10]

1. Thirteen states would retain CON.
2. Eight states would repeal CON.
3. Sixteen states would adopt rate-setting/review.
4. Six states would adopt hospital financial disclosure.

With some necessary cautions for the limitations of our data,[11] we used both simple and multinomial logistic regression in an effort to measure the extent to which the data support inferences concerning the effects of our fragmentation/cohesion alternatives on adoption or retention of spe-

---

[10] Because five states adopted rate-setting health policies and one state adopted financial disclosure prior to the time period under study, we have excluded them for purposes of our analysis. Louisiana, which never adopted CON nor moved to another policy, is also excluded. We do, however, report them in our initial data table and discuss the early rate setters in that section.

[11] The small number of data points in our model makes any statistical analysis weak. The marginal splits on our dependent variables tend to be outside desirable bivariate limits (i.e., no more extreme than 70–30). With only 43 cases, even 4 × 2 contingency tables produce some expected cell sizes less than five. We also expect that some misclassifications have occurred due to the crudeness of our measures We do not feel that from a theoretical perspective we can justify interpreting our independent variables as having interval qualities. Therefore in our regression we treat them as categorical.

cific health policies.[12] In addition, we employed chi square tests to examine patterns of association between the various configurations of fragmentation/cohesion and a reduced form of our expected policy outcomes.

In table 1 we present the array of states organized by degree of fragmentation within the hospital industry and, within that division, by level of business coalition development; for the high business development states we indicate the year in which the earliest autonomous coalitions were initiated; finally, we indicate the year in which the respective health policies were adopted (CON, having been adopted much earlier, is simply noted for its retention). Table 2 presents the results of our logit regression analyses.

The policy adoption patterns are highly congruent with our theoretical expectations. Eight of the 13 expected states (57%) retained CON. Four of eight expected states (50%) repealed CON (but only two others repealed among all other 35 states). Five of six expected states (83%) adopted financial disclosure. The results for rate setting were mixed and unclear (see discussion below).

Dichotomous and multinomial logit regression analyses with hospital fragmentation and business coalition development as the two independent variables produced similar results, both offering strong support for our model. A unified hospital industry promotes CON retention, while a fragmented hospital industry promotes CON repeal and passage of financial disclosure (multinomial model), though not significantly for the latter. Business coalition development significantly promotes the adoption of financial disclosure policy, significantly retards the retention of CON, and promotes the adoption of rate setting (multinomial model), though not significantly. We considered population growth (1970–80) as a possible alternative source of explanation. We found that it is significant in retarding CON retention and in promoting adoption of financial disclosure (only in the dichotomous model) and, as might be expected, somewhat reduces the significance of hospital fragmentation. This is consistent with our theoretic/historical argument that forprofit hospital development occurred especially in the Sunbelt, that is, in high population growth states, and that CON would be retained in nonfragmentation (and noncohesion) states. We regard population growth as a variable that precedes hospital fragmentation in time.

In addition, we reduced the four policy outcomes to two by grouping CON retention and hospital rate setting together as a "regulatory" policy outcome and CON repeal and hospital financial disclosure as a "competi-

[12] We used one-tailed tests of significance. For the multinomial model we compared the likelihood of adopting each of the three new policies, respectively, against that for retention of CON

TABLE 1

FRAGMENTATION/COHESION CHARACTERISTICS AND POLICY OUTCOMES

State	Forprofit Beds (%)	Business Coalition Score	Year Established	Retain CON Only	Repeal CON Only	Rate Setting	Financial Disclosure (without Rate Setting)
A. High Hospital Fragmentation/Low Business Coalition Development							
Texas	26.1	3			1985		
Alabama	16.1	3					
Georgia	13.4	3					
Kentucky	13.2	0		X			
Utah	12.7	0		X	1984		
West Virginia	11.7	0		X			
Arkansas	8.7	3			1987	1983	
New Mexico	7.4	0			1983		
B. High Hospital Fragmentation/High Business Coalition Development							
Nevada	43.6	4	1983				1987
Florida	28.3	4	1980			1984	
California	21.1	4	1981				1982
Tennessee	18.9	4	1975				1985
Lousiana*	16.5	4	1982				
Virginia*	14.8	4	1982				1978
Arizona	7.5	4	1981				1983
South Carolina	7.2	4	1981				1985

986

C. Low Hospital Fragmentation/Low Business Coalition Development

State			Established	X	Year
Mississippi	6.4	0		X	
Idaho	5.7	0	1982		
Hawaii	5.7	2			
Maryland*	5.4	0	1973	X	
Kansas	2.9	0	1985		
South Dakota	2.7	0		X	
Montana	1.2	0		X	
Iowa	.9	3		X	
North Dakota	.5	0		X	
Minnesota	.5	0			1984
Michigan	.2	0		X	
Wyoming	0.0	0			1985
Alaska	0.0	2		X	
Maine	0.0	2		X	1983

D. Low Hospital Fragmentation/High Business Coalition Development

State			Established	X	Year
Oregon	6.0	5	1982		1983
New York*	6.0	5	1979	X	1978
North Carolina	5.0	4	1985	X	
Washington*	4.9	4	1982		1973
Colorado	4.9	5	1984		1985
New Jersey*	3.9	5	1981		1978
Oklahoma	3.9	4	1981		
Missouri	3.5	4	1979	X	
Pennsylvania	3.3	5	1978	X	
Illinois	2.1	4	1980		
Indiana	1.4	4	1981		
Massachusetts	1.2	4	1981		1982
Nebraska	.6	4	1981	X	
Ohio	.1	4	1981	X	
Vermont	0.0	4	1982		1983
Delaware	0.0	4	1983	X	
Connecticut	0.0	4	1979		
New Hampshire	0.0	4	1982		1984
Rhode Island*	0.0	5	1984		1971
Wisconsin	0.0	4	1984		1983

NOTE.—Business coalition development is reflected in Business Coalition Score 3 or below indicates low development, 4 or 5 indicates high development. Year Established denotes the year in which the earliest autonomous business coalition with a development score of 4 or 5 was established.
* This state excluded from the statistical analysis due to the establishment of focal policy prior to the study period.

TABLE 2

EFFECTS OF HOSPITAL FRAGMENTATION AND BUSINESS COALITION DEVELOPMENT ON POLICY ADOPTION

	RETAIN CON		REPEAL CON		RATE SETTING		DISCLOSURE	
	b	SE	b	SE	b	SE	b	SE
A. Dichotomous Logistic Regression								
Hospital fragmentation	-1.5189**	8071	1.7047*	1.0044	.1346	9493	6455	.8285
Coalition development	-1 3456**	7046	-10 2185	55 4683	7624	9332	2 5498***	.8886
Constant	-7616	3954	5.9616	27 7343	-1 8558	4851	-9402	4428
B. Dichotomous Logistic Regression Including Population Growth								
Hospital fragmentation	-1 0566	.8886	1 6174	1.1364	8342	1.1744	-2884	9967
Coalition development	-1 5857**	7856	-11 7269	47 5416	8508	.9683	2 6936***	.9459
Population growth	-5.8412*	3.4472	8.0683	5 2747	-4 3887	4 4809	5.9559*	3 4207
Constant	2257	6762	-8 3678	23 8457	-1 0658	8574	-2.1839	8784

	REPEAL VS RETAIN CON		RATE SETTING VS RETAIN CON		DISCLOSURE VS RETAIN CON	
	b	SE	b	SE	b	SE
C. Multinomial Logistic Regression						
Hospital fragmentation	2 0660¹	1 1054	1 0311	1.2979	1.2488	.9953
Coalition development	-9.3107	65 688	1 4022	1.0310	2.5334**	9609
Constant	-1 4821	7821	-2 0170	8803	-2.1135	8791
D. Multinomial Logistic Regression including Population Growth						
Hospital fragmentation	2 17484*	1.24516	1.20395	1.1032	7771	1 09845
Coalition development	-10 391	57.6101	1 0487	1.45439	2 43651**	9605
Population growth	1 26824*	75255	59988	02539	49314	49316
Constant	-3.48921	1.68152	-2.09191	1 04468	-2 55197	1 02491

* $P < 10$, ** $P < 06$, *** $P < 01$

tion" policy outcome. We found a strong association between these outcomes and the four fragmentation/cohesion alternatives ($\chi^2 = 8.27$, $\alpha = .04$, $df = 3$).

While our quantitative analysis provides empirical support for our theoretical arguments, it does not account for the adoption of unexpected policy. Some of these outcomes may be at least partially congruent with or accounted for within our theoretical framework. Among states adopting unexpected alternatives, cohesion among business elites clearly fosters adoption of financial disclosure policy in that seven of the nine states not expected to adopt hospital financial disclosure policies (but did) met the criterion of high business coalition development. However, they did not exhibit high levels of forprofit hospital activity. Minnesota, one of the two states to adopt hospital financial disclosure in the absence of at least one highly developed, autonomous business coalition, is widely recognized as a leader in the development of cost-conscious managed care programs. No doubt the aggressive price negotiation posture of these organizations was an important factor in the adoption of that state's hospital financial disclosure policy. Wyoming's adoption of financial disclosure in the absence of at least one highly organized business coalition appears to support further the notion emerging from the analysis that, in some instances, powerful nonhealth elites possess the capability of moving state policy toward desired structural arrangements through less formal means, or through institutional mechanisms other than business coalitions.[13] We intend to investigate this further through our case study analyses.

In two of the three high-fragmentation states that maintained supply regulation (contrary to our expectations) the rate of increase of the forprofits (1972–86) exceeded 200% (the remaining state increased by 55%). This high rate of increase suggests that in those three states the regulatory agency was not an effective barrier to forprofit entry into the hospital market; and that these states may have retained supply regulation, not to restrict market entry, but precisely because the program did not, while at the same time it provided the appearance of doing so.

All six of the states moving toward open competition market relations (via repeal of CON and nonadoption of either of the other two focal policies) lie outside the industrial Northeast and Midwest. Only two of these states are among the 37 not expected to move toward market relations based upon open competition. Both of these states experienced high

[13] For example, our recent case study of health care politics in the state of Texas (see Rond 1990) documents the role of a "grass-roots coalition" of indigent-care advocates in attaining passage of policy initiatives detrimental to the interests of dominant business elites.

levels of state fiscal strain during the study period, and this strain may have inhibited the adoption of new rate setting or financial disclosure programs.

Five of the 11 states that adopted some form of mandatory hospital rate review did so before 1981. As such their policies predate the emergence of organized business coalitions and are excluded from our analysis. However, it should be noted that four of the five are located in the industrial Northeast—congruent with the historic regional dominance of the voluntary hospitals—are highly unionized, and have a long history of strong business influence and active state government. Accordingly, we would argue that in these instances the state most likely acted on behalf of cohesive (though, perhaps, not overtly organized) nonhealth elites in seeking to control the flow of revenues into the health sector.[14] With this interpretation, rate-setting states are better explained by our model.

In sum, our study indicates that during the period under examination there was strong movement away from market relations based principally upon supply regulation and toward other structural arrangements. Overall, the pattern of restructuring was consistent with the tenets of our theoretical framework. In general, failures to initiate policies congruent with state patterns of elite fragmentation/cohesion reflect nonadoption of policies rather than adoption of incongruent policies. As such, we view the current state of affairs as still in a process of transition away from federally fostered supply regulatory market relations toward arrangements more appropriate to the specific conditions of the individual states. The prevalence of high levels of fiscal strain during the first half of the decade of the 1980s apparently retarded the transition, making states reluctant to adopt new programs, and in two instances prompting what might be considered inappropriate repeal of hospital supply regulation. Our state case studies now underway indicate the importance of state fiscal strain in policy adoption, and our future analyses will include further investigation of its significance (see Rond 1990).

DISCUSSION

Recent declarations of a transformation in American medicine have been grounded primarily in observations and commentary on corporate entry into the American health care arena and in a presumption of changes in power and dominance within that arena (see Starr 1982; Derber 1984; Pollitt 1982; Whiteis and Salmon 1987; Brown 1986). Whether the cur-

[14] Alternatively, we might consider under what conditions it could be said that the state engages in autonomous action. We intend to investigate this further in our case study analyses.

rent declarations can be considered different from those of previous crises is dependent on empirical and theoretical analysis of the relation of current crises and changes to long-standing arrangements of power and patterns of dominance among elites in and outside the health care arena. Judgment of the significance of this "transformation" is just beginning; our work is one step toward completion of that task.

We can conclude from our work to this point that established patterns of elite dominance in health care have indeed come under attack, both from within and without the health care arena. Further, these challenges have surpassed the "symbolic politics" of 15 years ago. Where health care battles at that time were chiefly rhetorical and of little substance (Alford 1975), current policy battles are clearly substantive. Under the influence of business elites, state intervention has moved either to limit the degrees of freedom of health provider elites in favor of purchasers of care (i.e., to move toward price and product regulatory market relations) or to encourage a process of direct purchaser elite management over health provider elites (i.e., to move toward negotiated competitive market relations). Only in those states fostering open competitive market relations, where health provider elites potentially battle among themselves, and in those remaining states with predominantly supply regulatory market relations, do health provider elites potentially retain their historically dominant role.

Thus, apparently, given the major economic transformation that parallels these recent policy initiatives, traditional health provider elites—doctors and voluntary hospitals—will increasingly be displaced from their positions of dominance. But the long-term results of these policy initiatives are less clear. The implementation of new health care policies necessarily includes traditional health provider elites who have the most immediate control over the health institutions within which or for which the policies are implemented. While these elites are currently under attack, they will nonetheless have over time, or already have had, the opportunity to shape implementation in a direction consistent with their own material interests. Moreover, preliminary evidence suggests that business/corporate elites that moved to intervene directly in health policy-making may perceive the immediate crisis as having passed and may now be retreating from extensive or direct involvement in the implementation of policy (Brown and McLaughlin 1987). Unless the new political and economic power wielded by purchasers of care is used not just to sanction cost controls, but to effect permanent changes in market control and management, it is unlikely that a much broader transformation will occur. Absent such longer-term changes, health provider elites may be able to reassert their position of dominance in the provision of care, as they have in the past.

Nonetheless, recent state health policy has clearly fostered the continuing commodification of American health care and the development of new market relations. Given the continuing rise in health care costs and the human consequences of being "uninsured," the notion that health care should be treated as a commodity is likely to continue to evoke a new and different crisis of legitimacy.

REFERENCES

Alford, Robert R. 1975. *Health Care Politics*. Chicago University of Chicago Press.
Alford, Robert R., and Roger Freidland. 1985. *Powers of Theory*. Cambridge. Cambridge University Press.
Bergstrand, Curtis R. 1982. "Big Profit in Private Hospitals." *Social Policy* 13. 39–41.
Bergthold, Linda 1987. "Business and the Pushcart Vendors in an Age of Supermarkets." *International Journal of Health Services* 17:7–26
————.1990. *Purchasing Power in Health*. New Brunswick, N J : Rutgers University Press.
Bluestone, Barry, and Bennett Harrison. 1982. *The Deindustrialization of America: Plant Closings, Community Abandonment, and the Dismantling of Basic Industry*. New York: Basic Books
Brown, E. Richard. 1979. *Rockefeller Medicine Men: Capitalism and Medical Care in America*. Berkeley: University of California Press.
Brown, Lawrence D. 1986. "Introduction to a Decade of Transition." *Journal of Health Politics, Policy and Law* 11:569–83.
Brown, Lawrence D , and Catherine G. McLaughlin. 1987. " 'May the Third Force Be with You'. Community Programs for Affordable Health Care." Manuscript. Columbia University.
Burt, Ronald S. 1983. *Corporate Profits and Cooptation*. New York: Academic Press.
Davis, Mike. 1984 "The Political Economy of Late-Imperial America." *New Left Review* 143.6–38.
Derber, Charles. 1984 "Physicians and Their Sponsors. The New Medical Relations of Production." Pp. 217–54 in *Issues in the Political Economy of Health Care*, edited by John B. McKinlay. New York: Tavistock.
Ehrenreich, Barbara, and John Ehrenreich. 1970. *The American Health Empire: Power, Profits, and Politics*. New York: Random House
Ellwood, John William, ed 1982. *Reductions in U.S. Spending. How They Affect State and Local Governments*. New York: Random House.
Encyclopedia of Associations 1988 22d ed., vol. 1, pt. 2, pp. 1183, 1227.
Fine, Gary A. 1984. "Negotiated Orders and Organizational Cultures." Pp. 239–62 in *Annual Review of Sociology*, vol. 10. Edited by Ralph H. Turner and James F. Short. Palo Alto, Calif.: Annual Reviews.
Freidson, Eliot. 1970 *Professional Dominance*. New York. Atherton.
Hoff, John S., and Kenneth I Schaner 1982 "Government Policies Force Nonprofits to Go For-Profit." *Modern Healthcare* (June 1) pp 81–82, 84–85
Imershein, Allen W. 1977. "Organizational Change as a Paradigm Shift." *Sociological Quarterly* 18·33–43.
Imershein, Allen W., and Lance deHaven-Smith 1987. "The New Federalism in Florida." Pp 115–37 in *Reagan and the States*, edited by Richard P. Nathan, Fred C. Doolittle, and associates Princeton, N J. Princeton University Press.
Imershein, Allen W., and Philip C. Rond III. 1989. "Elite Fragmentation in Health

Care: Proprietary and Non-Proprietary Hospitals in Florida." *Social Science Quarterly* 70.53–71.

Intergovernmental Health Policy Project. 1985 *Major State Health Laws, 1984.* Washington, D.C.. Intergovernmental Health Policy Project.

King, Martha P., and Darrin Francis 1987. *1986 Health Care Legislation.* Denver: National Conference of State Legislatures.

Laumann, Edward O., and David Knoke. 1987. *The Organizational State.* Madison. University of Wisconsin Press.

Lee, Philip R., and Carroll L. Estes. 1983 "New Federalism and Health Policy " *ANNALS, AAPSS* 468:88–102.

Light, Donald. 1991. "The Restructuring of the American Health Care System " In *Health Politics and Policy,* 2d ed. Edited by Theodor Litman and Lee Robins. New York· Wiley.

Markowitz, G. E., and D N Rosner. 1973. "Doctors in Crisis A Study of the Use of Medical Education Reform to Establish Modern Professional Elitism in Medicine " *American Quarterly* 25:83–107.

McKinlay, John B , and John D. Stoeckle. 1988 "Corporatization and the Social Transformation of Doctoring " *International Journal of Health Services* 18.191–205

Mechanic, David. 1979 *Future Issues in Health Care. Social Policy and the Rationing of Health Services.* New York: Free Press.

National Conference of State Legislators (NCSL). 1985. *Hospital Cost Containment: A Legislator's Guide.* Denver. NCLS.

Pierce, Bob, and Ellen Hekman. 1984. *Health Care Cost Containment Legislation: 1983 Legislative Update for the Fifty States* Denver: National Conference of State Legislatures

Polchow, Michelle. 1986. *1985 Health Care Cost Containment Legislation.* Denver: National Conference of State Legislatures

Pollitt, Christopher 1982. "Corporate Rationalization of American Health Care. A Visitor's Appraisal " *Journal of Health Politics, Policy, and Law* 7:227–53.

Relman, Arnold S. 1980. "The New Medical Industrial Complex." *New England Journal of Medicine* 303, no 17:963–70.

Rond, Philip C., III. *Elite Fragmentation and Structural Change in Health Care.* 1990. Doctoral dissertation. Tallahassee· Florida State University

Salkever, D. S., and T. W Bice 1976. "The Impact of Certificate of Need Controls on Hospital Investment." *Milbank Memorial Fund Quarterly* 54:185–214.

Starr, Paul 1982 *The Social Transformation of American Medicine.* New York. Basic Books.

Stevens, Rosemary. 1971 *American Medicine and the Public Interest.* Philadelphia: Lippincott.

Useem, Michael 1983 "Business and Politics in the United States and United Kingdom." *Theory and Society* 12.281–308

————. 1984. *The Inner Circle: Large Corporations and the Rise of Business Political Activity in the U.S. and U.K.* New York. Oxford University Press.

Whiteis, David, and J Warren Salmon. 1987 "The Proprietarization of Health Care and the Underdevelopment of the Public Sector." *International Journal of Health Services* 17.47–64

Dynamics of Professional Control: Internal Coalitions and Crossprofessional Boundaries[1]

Sydney A. Halpern
University of Illinois at Chicago

Professions experience varying degrees of success in establishing jurisdictional control over neighboring occupations. This article uses comparative historical techniques to explore why some American medical specialties were more successful than others in dominating adjacent professions. Three major traditions in the sociology of professions are unable to fully account for such differences. This article offers an alternative explanation. Specialties best able to subordinate ancillary workers are those receiving support from established segments within medicine. Thus, relations between segments within a dominant profession powerfully affect its boundaries with other occupations.

An emerging focus in the sociology of professions concerns interrelations between legitimate professions and their sometimes competing, often subordinate, neighboring occupations. Examples of interprofessional conflict where one contender prevails are found in many occupations. American accountants established superiority over bookkeepers early in the 20th century. Contemporary lawyers quelled encroachment from notaries, magistrates, real estate agents, and (now defunct) conveyancers (Abbott 1988). Scholarship suggests that contests for control are especially common among the very large number of professions involved with health services. Medicine has sought to contain many of these professions. But like other groups seeking occupational dominance, physicians have won greater control over some of their neighbors than over others. The following analysis explores the origins of such differences in the effectiveness of professional dominance.

This article examines variation in American medicine's control over four health occupations: radiologic technology, laboratory technology, physical therapy, and nurse anesthesia. During the 1920s and 1930s, groups of physicians moved to curtail encroachment from these ancillary professions by circumscribing the latter's occupational territory. Medical segments experienced varying degrees of success in excluding technical workers from tasks that doctors claimed as wholly their own. One group of physicians failed to check encroachment from nurse anesthesia by the late 1930s but strengthened its position during the 1940s and 1950s. Another group contained physical therapy's domain at the onset of World War II but lost its hold in the two decades that followed. Still others secured a comparatively high level of jurisdictional control over radiologic and laboratory technology by 1940 and retained it through the early 1960s. The question addressed here is what accounts for varying degrees of success in winning and maintaining (or enhancing) professional dominance, which is defined here as control over the jurisdiction of a neighboring occupation.

Sociological literature offers three approaches for explaining differential outcomes in contests for occupational control. The first approach suggests that attributes of the profession being subordinated may influence the effectiveness of occupational dominance. It derives from the work of professionalization theorists who argue that an unfavorable gender or social class composition or an insufficiently complex body of knowledge can impede an occupation's efforts to achieve full professional standing (Etzioni 1969; Ritzer 1977; Simpson and Simpson 1969). The second approach links occupational control to the extent of the dominant profession's hold over subordinates' occupational structures: their professional associations, certification systems, training programs, and accrediting bodies (Brown 1973; Freidson 1970). This perspective is grounded in professional power theory, which sees occupations competing for zero-sum quantities of authority and which elaborates the influence of political and institutional processes on the outcome of interprofessional struggles.

The third approach is Andrew Abbott's professional systems model. Abbott (1988) sees jurisdictional disputes—struggles for control over arenas of work—as the key events among professions. These conflicts occur when social forces create or extinguish work domains and when one profession in the system abandons or moves to encroach upon an occupied arena. Most jurisdictional struggles, he contends, take place in three locales: the workplace, public opinion, and legal and administrative rule. Conflicts result in a variety of settlements ranging from full control by a single profession, to split jurisdiction (e.g., by client type), to the subordination of one profession by another. While acknowledging that a number of factors affect the outcome of these struggles, Abbott emphasizes cogni-

tive factors, particularly the participants' ability to maintain an optimal level of abstraction. From this approach, the outcome of boundary disputes rests on the contenders' relative cognitive strengths.[2]

The data examined here point to limitations in each of these models. This paper presents original, comparative historical research on sources of variation in occupational control. It treats professional dominance as an ordinal concept and examines occupational subordination at two points in time: during formative years when professional boundaries were first established and a quarter century later when the jurisdictions' durability to challenges would become apparent. The analysis supports general predictions of both the power and systems perspectives by identifying a combination of cognitive and institutional factors affecting the outcome of contests for occupational control. However, these factors come to play in very different ways than is predicted by existing models.

None of the three theories addresses a crucial dynamic shaping the outcome of jurisdictional disputes: the influence of relations between segments within a profession on its boundaries with neighboring occupations—the impact of *intra*professional relations on *inter*professional boundaries. Sociologists in the tradition of the Chicago school were first to comment on the theoretical significance of professional segments (Bucher and Strauss 1961).[3] Other scholars examine the origins and character of segments within medicine (Halpern 1988; Rosen 1944; Rothstein 1972; Stevens 1971). These occupational units have interests, cultures, and formal associations distinct from those of medicine as a whole (Bucher and Strauss 1961). Some are organized into formally constituted specialties that function like "professions within a profession" (Halpern 1988). Because organized segments are similar to professions, one can apply Abbott's framework both to relations between specialties within a single profession and to relations between professions. But neither his nor the other two models adequately clarifies interdependencies between the two levels of jurisdictional units.

---

[2] "Cognitive components dominate the structuring of jurisdictions" as well as "the conduct of jurisdictional contests "(Abbott 1988, p. 98). While Abbott offers the most fully articulated available model of interprofessional relations, elements of his framework can be found in earlier scholarship. Goode (1969) addresses interrelationships among evolving professions on a theoretical level Empirical treatments of interprofessional boundary disputes preceding Abbott's include those by Albrecht and Levy (1982), Begun (1987), Goode (1960), and Kronus (1976). Abbott's emphasis on occupational jurisdiction has antecedents in research on craft labor markets (Jackson 1984). The importance of cognitive factors is a longstanding theme in the sociology of professions (Carr-Saunders and Wilson 1933)

[3] Consistent with Bucher and Strauss (1961), I use *segment* to refer to occupational groupings whose members share common work activities, values, identities, and missions.

Intraprofessional dynamics matter in part because specialties have different interests and levels of involvement in boundary settlements with particular neighbors. For the majority of occupations subordinated by medicine—general nursing being the most noteworthy exception—professional dominance is initiated less by medicine as a whole than by a formally organized medical specialty. The specialties that subordinate the ancillary professions examined here (radiologic technology, laboratory technology, physical therapy, and nurse anesthesia) are, respectively, radiology, clinical pathology, physiatry (also known as physical medicine and rehabilitation), and anesthesiology. Following World War I, each of these specialties—still themselves fledgling occupational groups—spearheaded efforts to circumscribe the terrain of its ancillary work force. The American Medical Association (AMA) was involved as well; by the mid-1930s, that association was overseeing the accreditation of training programs for three of the ancillary professions. But specialty associations, not the AMA, launched and directed efforts to delimit the jurisdictions of ancillary workers.

If one specialty typically initiates professional dominance, it needs the cooperation of other segments to effectively subordinate its ancillary work force. Data presented here show that an emerging specialty's ability to control its technical personnel rests on the willingness of established segments to acknowledge the former as a distinct intraprofessional jurisdiction. Subordination of a neighboring occupation thus begins with collaboration among segments within the dominant profession.

The remainder of this article is organized into four major sections. The first compares specialities' success at circumscribing the jurisdictions of ancillary workers. It documents variations in professional dominance to be accounted for in the subsequent analysis. Section 2 evaluates specific explanations for degrees of subordination derived from sociological literature. Section 3 develops the article's central arguments concerning intraprofessional determinants of professional dominance. In doing so, it explores the influence of cognitive and institutional contingencies on policies formulated by established professional segments. Section 4 elaborates implications of the historical analysis for theories of interprofessional relations, particularly for Abbott's systems model.

GRADIENTS OF JURISDICTIONAL SUPERIORITY

For the purpose of this paper, jurisdictional control is achieved when a medical specialty succeeds in imposing a division of labor whereby technical personnel are excluded from activities that the physician group claims as strictly its own. Two empirical indicators attest to the strength of jurisdictional control: the prevalence of physician complaints about

boundary infringement (dissatisfaction implying poor control) and the extent of hierarchical task differentiation between specialists and ancillary workers. Data on these indicators from contemporary medical literatures reveal discernible patterns in the effectiveness of jurisdictional subordination.

During the 1920s and early 1930s, all four medical specialties complained of encroachment from ancillary workers. Radiologists decried the tendency for x-ray technicians to go beyond the legitimate function of producing plates by interpreting the results and offering medical diagnoses (Desjardins 1931; LaField 1923; Pancoast 1933; Perry 1933). Pathologists objected to violations by laboratory workers who not only processed samples but also commented on the medical implications of test results (Burdick 1924; Hillkowitz 1924; Kolmer 1925). Both radiologists and pathologists inveighed against lay-administered laboratories (so-called commercial labs) that competed with physician-run, hospital-based services (*Archives of Physical Therapy, X-Ray, and Radium* 1926; Desjardins 1931, p. 1751; Hillkowitz 1924; White 1979). Physiatrists—insisting that technical personnel limit themselves to routine, physician-ordered treatment—rebuked physical therapists for prescribing remedies and engaging in "medical piracy" by operating independent practices (*Archives of Physical Therapy, X-Ray, and Radium* 1927, 1934; Kovacs 1934, p. 686; also Smith 1936, p. 620). Medical anesthetists took their grievances beyond the confines of professional forums. Between the two world wars, regional anesthesiology societies sought—and failed to obtain—the legal prohibition of nurse-anesthesia practice (Betcher et al. 1955, p. 769; Thatcher 1953, pp. 110–24, 132–52).

The prevalence of encroachment complaints during interwar years can be attributed to the absence of established occupational boundaries. Both the specialties and their technical occupations were in a period of institutionalization. Occupational associations for the four ancillary professions first appeared between 1920 and the mid-1930s. All four specialties had established national professional societies by the early 1920s. But these segments did not assume their mature form until the creation of certifying boards, which began to proliferate in the mid-1930s. In the 1920s and early 1930s, the medical specialties were still struggling to delineate and secure their occupational territory.

Outcomes as of 1940

By World War II, initial boundaries across the health professions stabilized, and specialties began to differ notably in their degree of jurisdictional superiority. Radiology and clinical pathology achieved high levels of jurisdictional control. Physiatry was moderately successful at subordi-

nating ancillary workers, while anesthesiology was unable to secure its boundaries (see table 1 below for a summary of the specialties' evolving jurisdictional status).

Statements by contemporary radiologists attest to that specialty's success in eliminating competition from ancillary workers. As early as the mid-1930s, the head of a major roentgenology society remarked in his presidential address that "lay technicians, who render medical reports, have almost passed" (Murphy 1933, p. 719). Other leaders echoed this assessment (Perry 1933, p. 75). Meanwhile, commercial x-ray labs disappeared as hospital-based services became the industry standard. In radiology, the problem of encroachment, including competition from lay-administered laboratories, was "virtually eradicated" by 1940 (Stevens 1971, p. 230).

Pathology obtained a favorable boundary settlement even though it never entirely eliminated commercial laboratories, which continued to operate during the middle decades of the century and constituted a small segment of the laboratory industry (White 1979). However, specialists' principal competitors in the commercial sector were not ancillary workers but rather Ph.D.-trained microbiologists and biochemists who insisted on their own competency to oversee laboratory services (Magath 1939). The hospital sector of the industry was by far the more dominant one, and here technicians deferred to pathologists' jurisdictional claims (White 1979). In the mid-1930s, the president of the American Society of Clinical Pathologists (ASCP) declared: "the important *auxiliary* position of the laboratory worker . . . is now firmly established" (Simpson 1933, p. 13; emphasis added). By the onset of World War II, radiologists and clinical pathologists had gone far in buttressing a hierarchical division of labor: physicians delegated to ancillary workers the routine tasks of collecting and processing laboratory data; with increasing regularity, specialists alone handled clinical interpretation and diagnosis.

TABLE 1

LEVELS OF SPECIALTY CONTROL OVER THE JURISDICTION OF ANCILLARY WORKERS: LATE 1930s AND EARLY 1960s

	LEVEL OF JURISDICTIONAL CONTROL	
SPECIALTY	Late 1930s*	Early 1960s*
Anesthesiology	Low	Moderate
Clinical pathology..	High	High
Physiatry	Moderate	Low
Radiology	High	High

* The values below are time-specific ordinal rankings

Physiatrists also managed to curtail boundary infractions by ancillary workers. During the late 1930s, complaints about physical therapists prescribing and maintaining independent practices disappeared from the medical journals. The empirical record does point to considerable task overlap between physiatrists and physical therapists—greater than that between radiologists or pathologists and their technicians. Physiatrists and physical therapists both cared for patients and, in doing so, employed many of the same treatment modalities. (For this reason, I rank physiatrists' level of jurisdictional control as "moderate" in table 1.) Still, specialists were satisfied that physical therapists were no longer prescribing, and this eliminated the source of earlier encroachment complaints.

In the field of anesthesia, physician dissatisfaction continued unabated, although jurisdictional struggles took a different form than in previous decades. Anesthesiologists abandoned efforts to have nurse anesthesia declared illegal after meeting defeat in a landmark case during the mid-1930s (Thatcher 1953). The specialty failed to win favorable judicial rulings because it lacked the support of a key intraprofessional segment. "The American College of Surgeons came out against any legal restrictions which would prohibit trained nurses from giving anesthetics" (Stevens 1971, p. 239, n. 50). But anesthesiology's goal remained the curtailment of nurse anesthesia practice. The specialty pursued this aim in the 1940s by attempting to close the ancillary professions' training programs. The American Society of Anesthesiologists (ASA) censured members who participated in the education of nurse anesthetists, and specialty leaders strongly discouraged nursing schools from sponsoring anesthesia training (Beecher 1948, pp. 116; Betcher 1982, pp. 203; Bankert 1989, pp. 134–35; Gunn 1983, pp. 162–63). The medical segment eventually succeeded in removing nurse-anesthesia training from both university settings and schools of nursing (Freeark 1981, p. 566; Gunn 1983, p. 162). But the problem of overlapping jurisdictions remained. During the 1930s, no task hierarchy differentiated the professional roles of anesthesiologists and nurse anesthetists. With the exception of small numbers of medical anesthetists conducting research, members of the two occupations were carrying out virtually the same professional activities.

Durability through the Early 1960s

Jurisdictional settlements are subject to change over time, even those embedded in established institutional arrangements. Two of the specialties examined here experienced noteworthy alterations in their level of jurisdictional control during the aftermath of World War II; in the other two cases, professional boundaries remained quite stable. Radiology and clinical pathology sustained high levels of professional dominance be-

tween the late 1930s and the early 1960s. Both specialties continued to evolve in the 1940s and 1950s. New types of procedures and equipment became standard technologies within x-ray and clinical laboratories. Private health insurance covering diagnostic tests become increasingly widespread, and the volume of services provided by radiology and pathology units rose dramatically (Brecher and Brecher 1969; Rothstein 1979). Specialists mobilized to replace salaried positions with agreements for reimbursement on a fee-for-service basis (Stevens 1971). But the historical record gives no evidence of rekindled physician dissatisfaction over technicians' professional activities. Indeed the character of ongoing change was more likely to augment than diminish specialists' jurisdictional superiority.

Physiatry, however, failed to retain its previous level of jurisdictional dominance. In the mid-1950s, encroachment from ancillary workers re-emerged as a central theme in the editorial columns of the major physiatry journal. Physical therapists were demonstrating, in the words of one medical specialty leader, "a tremendous upsurge of professionalistic spirit" (Rose 1959, p. 3). Physiatrists' complaints echoed those made in earlier decades: physical therapists were avoiding medical supervision and pursuing independent practice (*Archives of Physical Medicine and Rehabilitation* 1953, 1954, 1955). The unsupervised physical therapist was, in effect, prescribing medical treatment. Official policy of the American Physical Therapy Association (APTA) supported private practice if the therapists worked closely with referring physicians. But a more militant segment within the ancillary profession was questioning the necessity of direct supervision by doctors. In the early 1960s, two splinter organizations formed outside the APTA for the purpose of promoting independent practice (Gritzer and Arluke 1985). Physiatry was losing control over what it designated as core medical tasks in the face of boundary expansion by ancillary workers.

Meanwhile, anesthesiology was enhancing its jurisdictional position. Modification in the specialty's stance toward nurse anesthesia was one sign of change. Leading anesthesiologists advocated that the specialty abandon its efforts to eliminate the ancillary profession and instead treat nurse anesthesia as an auxiliary work force (Beecher 1962; Bendixen 1967; Dripps 1962). During the early 1960s, the ASA reassessed its relations with the American Association of Nurse Anesthetists (AANA) and rescinded its policy of censoring anesthesiologists involved with the training of nurse anesthetists (Betcher 1982; Gunn 1975, p. 132). Even more significant was evidence of a new cross-professional division of labor. The postwar years saw increasing differentiation in the tasks performed by anesthesiologists and nurse anesthetists. Medical specialists were taking over the more complex spinal, intravenous, and nerve block proce-

dures (Sanders 1945)—performed with increasing frequency after World War II—and leaving the simpler general (inhalation) anesthesias to nursing personnel. By the 1970s, general anesthesias constituted more than 85% of nurse anesthetists' practice. Only a minority of nurses performed IVs, epidurals, or other blocks; and among those who did, the complex procedures constituted a tiny fraction of their overall practice (Goff 1977, pp. 150–52). Task overlap still existed. The two groups competed for general anesthesias and quite a few nurses were performing spinal procedures. (For this reason, I rank anesthesiology's jurisdictional control during the early 1960s as "moderate" in table 1.) Nonetheless, medical specialists had made discernible progress in establishing a hierarchical division of labor within anesthesia services.

In summary, radiologists and clinical pathologists secured high levels of jurisdictional control by World War II and retained them through the early 1960s. Physiatrists were moderately effective at dominating technical workers during the late 1930s but lost their hold in the two decades that followed. Anesthesiologists failed to subordinate nurse anesthetists before the war but appreciably increased their jurisdictional control by 1960—albeit not to the levels sustained by radiology or clinical pathology.

EXPLANATIONS DERIVED FROM EXISTING THEORIES

Each of three major traditions in the sociology of professions offers specific predictions about sources of variations in professional dominance. Professionalization theorists point to characteristics of the subordinated occupations as factors contributing to gradients of jurisdictional control. This framework categorizes ancillary health fields as semiprofessions: occupations that seek to establish themselves as full-scale professions but fail because of attributes like their predominantly female composition, the modest socioeconomic status of members, or the absence of a systematic body of theory (Etzioni 1969; Ritzer 1977; Simpson and Simpson 1969). This approach invites focused comparisons of the four technical professions that examine how occupational traits affect levels of professional dominance. More specifically, it predicts that the most autonomous and least subordinated of the ancillary professions would be those with the lowest proportion of women, the heaviest recruitment from the middle and upper socioeconomic strata, and the most complex bodies of knowledge. Historical data fail to substantiate these predictions.

The gender composition of all four technical professions was largely female during the middle third of the century. Nonetheless, some variation across fields is evident. Both radiology and medical technology had a sizable minority of men. Between 25%–30% of registered x-ray techni-

cians were men during the early 1950s (Greene 1954, pp. 82, 85). The proportion was probably larger in the prewar years; discussions of the original x-ray technology work force repeatedly refer to male recruits (Perry 1930, p. 35; Greene 1954, p. 77). Twenty percent of the early medical technology registrants were men, and this proportion remained quite stable over time (Montgomery 1970, p. 437; White 1979, p. 64). Physical therapy and nurse anesthesia had virtually no men within their ranks through 1960. Discussions of the physical therapy work force of the pre- and postwar decades mention only women (Hazenhyer 1946; American Registry of Physical Therapy Technicians 1950). Nurse anesthesia admitted no men into its national association until after World War II (Thatcher 1953, p. 256). Thus, contrary to prediction, the most autonomous ancillary fields (nurse anesthesia in 1940 and physical therapy in 1960) had fewer men than the most thoroughly subordinated fields (x-ray and laboratory technology).[4]

The best available indicator of the socioeconomic status of occupational recruits is education prerequisites for admission into professional training schools. In the late 1930s and 1940s, x-ray technology required only a high school diploma for entrance to its programs; some schools even waived completion of secondary education (AMA 1947, p. 1144; 1948, pp. 1456–59). Admission standards for nurse anesthesia, medical technology, and physical therapy were roughly equivalent. Nurse anesthesia required a degree in nursing—most often obtained from a two- or three-year hospital-based diploma program. Laboratory technology and physical therapy required two years of college, or a nursing degree, or (in the case of physical therapy) graduation from a school of physical education (AMA 1936, pp. 677, 682; 1942, p. 1135). Again, the data are inconsistent with expectations derived from theory. The ancillary profession with most autonomy in 1940 (nurse anesthesia) was apparently not recruiting more of its members from the middle and upper socioeconomic strata than were two of the highly dominated occupations (medical technology and physical therapy).

Data on knowledge complexity also fail to corroborate theoretical prediction. The best available measure of cognitive complexity is duration

[4] The gender composition of the medical specialties may have affected jurisdictional settlements. Calmes (1984) and Stevens (1971, p. 240) point out that, during the interwar years, anesthesiology had a high proportion of women within its ranks when compared to medicine at large. Consistent with the prediction of Cook, Moris, and Kinne (1982, p. 21), this might have contributed to anesthesiology's failure to establish interprofessional dominance in the late 1930s. However, scholarship provides no data on changes in the proportion of women in anesthesiology during the middle decades of the century, so it is not possible to evaluate whether shifts in gender composition may have contributed to later improvements in the specialty's jurisdictional status.

of professional training. In the late 1930s and 1940s, the majority of programs in x-ray technology, laboratory technology, and physical therapy were one year in duration (American Society of Radiographers 1932; AMA 1936, pp. 678, 680–81; 1942, pp. 1135–36; 1947, p. 1144). In laboratory technology and physical therapy, this year of training could be obtained only after two years of college or nursing school had been successfully completed. The American Association of Nurse Anesthetists (AANA) advocated year-long programs but required only six months of training in the late 1930s—in addition to the two years of nursing school required for entrance to the program (Thatcher 1953, pp. 227, 234). Thus, the most autonomous occupation in 1940 (nurse anesthesia) had a course of training no longer—and in many cases, shorter—than that of the more strongly subordinated occupations. Overall, then, explanations derived from professionalization theory fail to account for differences in the extent of the ancillary occupations' subordination.

The specific prediction drawn from Abbott's systems model fares little better in light of the empirical record. This framework suggests that the outcome of struggles for jurisdictional control rests upon the contending professions' relative intellectual strengths. The next section of the paper—which explores conditions shaping intraprofessional policy on jurisdictional settlements—shows that the actors adjudicating cross-professional boundaries were, indeed, attentive to cognitive factors. But decision makers were concerned not with the contending specialty's cognitive superiority to technicians but rather with the merits of its claim to expertise not shared by other doctors.

Historical data do provide strong support for the explanation offered by professional power theorists. Scholars adopting this framework link professional dominance to physicians' control over ancillary workers' occupational institutions (Brown 1973; Freidson 1970). Medical specialties began creating professional structures for technicians during the early 1920s. By the mid-1930s, a number of specialties were directing training programs, certifying boards, and accrediting systems for ancillary workers. Table 2 summarizes historical patterns of specialty control over technicians' professional institutions. Together with table 1, it shows that structural dominance and jurisdictional control often coincide.

For radiology and clinical pathology, physician management of technicians' occupational institutions accompanied high levels of jurisdictional control in both the late 1930s and the early 1960s. Radiologists stimulated the creation of the American Society of X-Ray Technicians (ASXT) in 1920. They directed educational programs for ancillary workers, inaugurated a technician registry in 1922, and began accrediting training schools in 1933. Clinical pathologists played a similar role with their ancillary labor force. Specialists established schools for laboratory workers, a tech-

TABLE 2

PATTERNS OF SPECIALTY DOMINANCE OVER THE OCCUPATIONAL STRUCTURES OF
ANCILLARY WORKERS· MID-1930S AND LATE 1950S

| | DOES THE SPECIALTY ADMINISTER TECHNICIANS' OCCUPATIONAL INSTITUTIONS? | |
SPECIALTY	Mid-1930s	Late 1950s
Anesthesiology	No	No
Clinical pathology	Yes	Yes
Physiatry	Yes	No
Radiology	Yes	Yes

nician registry in 1928, an occupational association in 1932, and an ac-
crediting body in 1935. Over time—particularly after World War II—
leaders of the ancillary professions pushed for greater control over their
own occupational structures. These efforts produced some modest adjust-
ments. Laboratory technicians won representation on their certifying
board. X-ray technicians became registry cosponsors. But specialists
loosened their reins only sparingly. Laboratory technology attempted but
failed to gain separate incorporation of its certifying board in the 1950s.
As of 1960, radiology and clinical pathology still dominated their subordi-
nates' occupational institutions.

Physiatry attained structural dominance and jurisdictional superiority
by the late 1930s. It lost both before 1960. The specialty helped establish
a national association for physical therapists in 1921. It oversaw a registry
for ancillary workers from the mid-1930s, and physical therapy training
schools of the prewar period all listed physician directors (AMA 1936,
pp. 677–78). Unlike radiology and clinical pathology, physical medicine
secured control over technicians' occupational structures with the help of
the AMA, which pressured the APTA to accept physiatrist-administered
training and credentialing (Hazenhyer 1946.) Also, unlike the other two
specialties, physiatry failed to sustain its structural dominance after
World War II, when ancillary workers moved for control over their own
professional institutions. I argue below that lack of support from key
intraprofessional segments undermined physiatry's hold on technicians'
occupational structures. But, for the moment, suffice it to say that physi-
cal therapists initiated a campaign for state licensure in the 1940s and
succeeded in obtaining it in most states by the early 1950s (Robinson
1956). The physician-controlled registry eventually fell into disuse
(Gritzer and Arluke 1985). Meanwhile the APTA sought and eventually
won control over the accreditation of its training programs (Evans 1978).

Strong patterns of covariation between structural dominance and juris-

dictional control notwithstanding, the professional power framework leaves a number of crucial questions unanswered.[5] Most compellingly, what accounts for anesthesiology's improved jurisdictional status during the postwar years? At no point did this specialty manage nurse anesthetists' occupational structures. Anesthesiology improved its jurisdictional position despite continued lack of control over technicians' professional institutions. Its success cannot be attributed to shifts in public opinion or judicial rule. Between World War II and 1960, neither anesthesiology or nurse anesthesia sought to resolve their disputes in public arenas. Nor did the federal or state governments attempt to intervene. These observations suggest that still other social processes were influencing boundaries between medical specialties and ancillary professions.

INTRAPROFESSIONAL PROCESSES SHAPING
BOUNDARY SETTLEMENTS

Existing theories fail to consider the impact of relations among segments within a profession on its boundaries with other occupations. In order to achieve jurisdictional control, a professional segment initiating the subordination of its ancillary work force needs support from established segments. Nowhere was broader professional support more vital to medical specialties than in the formulation of hospital standardization policies. Medicine had consolidated authority to shape the division of labor within hospitals by World War I, and it sustained this authority through the early 1960s (Rosenberg 1979, 1987; Starr 1982). Among the profession's internal segments, surgery was the key player in the ordering of hospital services. American hospitals became increasingly oriented toward surgical care as the institutions modernized in the decades before and after the turn of the century. By the early 1900s, the surgeon was "kingpin of hospital operations and prestige" (Stevens 1989, p. 180). It was the American College of Surgeons (ACS) that spearheaded and administered

[5] Among these questions is why radiology, clinical pathology, and physiatry established control over subordinates' occupational structures during the 1920s and 1930s while anesthesiology did not. Another is why physiatry lost its hold over technicians' professional institutions following World War II while radiology and clinical pathology retained theirs. A complete explanation for these patterns can be found in Halpern (1991) Suffice it to say here that three interrelated factors strongly influenced whether a medical segment sought and won structural dominance during the interwar period and sustained it over time (1) the extent of resources available to the specialty, (2) the segment's intraprofessional standing, and (3) the timing of the segment's development as a professionally self-conscious unit relative to the development of its ancillary labor force. A specialty could more easily establish and maintain structural dominance if it initiated administrative control over subordinates before technicians moved to establish their own (autonomously managed) occupational institutions.

the profession's hospital standardization program, beginning in the second decade of the century.[6] A Joint Commission for the Accreditation of Hospitals (JCAH) took over hospital inspection and accreditation activities in 1952. But in the preceding decades, the surgeons' national professional association—in consultation with an array of contemporary medical societies—set industrywide standards for hospital organization. The ACS approval conferred status, and hospitals with sufficient resources to conform to its requirements—at first the larger institutions and those in urban settings—endeavored to do so. While compliance with ACS policies was voluntary, their impact on the organization of medical work was substantial.

Hospital standardization guidelines affected boundaries between specialties and ancillary professions in two distinct ways. First, by requiring some but not other services to be hospital based, ACS policy influenced the extent to which types of work were centralized within hospitals or dispersed in office settings. A specialty could more readily establish jurisdictional control if subordinates were centralized in hospitals, where other work rules applied. Second, by requiring some but not other hospital departments to be headed by dedicated specialists, ACS guidelines designated which physicians would be supervising an ancillary work force. A specialty was best able to subordinate technicians if it received an ACS mandate for exclusive supervision of its ancillary work force. Specialties with exclusive supervision could enforce their notion of appropriate boundary settlement because they controlled both the allocation of tasks within work settings and the labor market for ancillary workers. As supervisors managing the hiring and firing of subordinates, specialists wielded the implicit threat—seldom needing enforcement—that they could deny jobs to technicians who committed jurisdictional violations.

If exclusive supervision was important for a subordinating specialty, it also had consequences for the broader community of doctors. The arrangement restricted prerogatives of other practitioners because it interposed a specialist between the technician and the referring physician. In the early decades of the century, many doctors were quite willing to bypass medical consultants (radiologists, pathologists, physiatrists, and anesthesiologists) and go directly to technicians for services. Surgeons welcomed nurse anesthetists into the operating room. Internists were

---

[6] The ACS's interest in hospital reform was an outgrowth of its concern with standards of surgical practice (Stevens 1971, p. 87). The AMA also considered assuming the leadership of hospital standardization but it had already invested considerable resources in its program to inspect and upgrade medical colleges The Carnegie Foundation chose to give a series of grants for hospital standardization to the ACS rather than the AMA (Stevens 1971, pp. 91, 119; Davis 1973, pp. 175–76).

happy to receive test results directly from x-ray and laboratory technicians (Stevens 1971, p. 226). Orthopedists routinely sidestepped physiatrists when referring patients for physical therapy. When a segment obtained exclusive supervision, referring clinicians no longer had access to technical services except through a dedicated specialist. Awarding a specialty exclusive supervision of its ancillary workforce was tantamount to granting it a protected intraprofessional jurisdiction. At issue was both the division of labor among medicine's internal segments and the profession's boundary with a neighboring occupation.

Table 3 illustrates historical patterns of ACS support for exclusive supervision over technical personnel. The empirical record shows a strong relationship between jurisdictional control and hospital standardization policies. For three of the specialties examined here, jurisdictional superiority postdated favorable standardization guidelines. Radiology and clinical pathology—both with high levels of jurisdictional control in the late 1930s—received early endorsement of exclusive supervision. During the mid-1920s, ACS leaders stipulated that, wherever possible, dedicated specialists should head x-ray and pathology departments (American College of Surgeons 1925, pp. 12–13). Furthermore, the very first guidelines required all hospitals to offer laboratory and radiology services, thus encouraging the centralization of work in settings where ACS supervisory policies would apply. Historical sources do not reveal precisely what portion of hospitals had specialists directing their x-ray and pathology units during the late 1920s and 1930s. Identifying specialists of this period was difficult; the radiology and pathology certifying boards were not established until the mid-1930s. Nonetheless, contemporary medical literature suggests that hospitals were seeking and rapidly acquiring radiologist and clinical pathologists as department heads;

TABLE 3

PATTERNS OF INTRAPROFESSIONAL SUPPORT FOR EXCLUSIVE SUPERVISION OF
ANCILLARY WORKERS: MID-1930s AND LATE 1950s

| SPECIALTY | DOES THE SPECIALTY RECEIVE SUPPORT FOR THE EXCLUSIVE SUPERVISION OF TECHNICIANS FROM HOSPITAL STANDARDIZATION GUIDELINES? | |
	Mid-1930s	Late 1950s
Anesthesiology	No	Yes
Clinical pathology	Yes	Yes
Physiatry	No	No
Radiology	Yes	Yes

where this was not the case, a shortage of specialists was the primary obstacle.

Anesthesiology also improved its jurisdictional position only after receiving support for exclusive supervision. But these events took place a decade and a half later in anesthesiology's development than in the history of radiology and clinical pathology. Standardization guidelines of the 1920s were silent on the management of anesthesia services. During the early 1930s, ACS publications stated a preference for direction by a physician with some training in anesthetics (American College of Surgeons 1933, pp. 31, 53). This standard was sufficiently broad to include a great many surgeons. It was not until the late 1930s that guidelines specified anesthesiologists as department heads (American College of Surgeons 1938, p. 43). In the middle third of the century, American hospitals did substantially alter the management of their anesthesia units. Teaching institutions began appointing anesthesiologists as service directors in the 1930s (Adriani 1972; Miller 1933; Volpitto and Vandam 1982). Widespread diffusion of anesthesiologist-headed units occurred in the late 1940s and 1950s. During the 1960s, anesthesiology leaders still complained about hospitals where medical direction of nurse anesthetists was inadequate (Ballinger 1962, p. 137). Nonetheless "the right of an anesthesiologist to supervise and direct the activities of nurse-anesthetists [had] been clearly established" (Bunker 1972, p. 58). Thus, for three of the four specialties, intraprofessional support for favorable supervisory arrangements preceded improved jurisdictional status.

One specialty, however, does not fit the prevailing pattern. Physiatry wielded moderately high levels of jurisdictional control in the late 1930s despite an absence of support for exclusive supervision. At no point did ACS policy specify physiatrist heads for physical therapy units nor did it even require that hospitals have physical therapy departments. With little impetus for centralization, physical therapy services would disperse outside hospitals where the division of labor between medical specialists and ancillary workers was difficult to regulate. With no mandate for exclusive supervision, physiatrists failed to secure the direction of physical therapy units located in hospital settings. Orthopedic surgeons, radiologists, and even technicians managed hospital-based physical therapy services during the middle third of the century (Kovacs 1934, p. 685; Hibben 1930, p. 467; Gritzer and Arluke 1985, p. 68).

The findings summarized in tables 2 and 3 invite speculation about the relative importance of exclusive supervision and structural dominance to the achievement of favorable jurisdictional settlements. The example of anesthesiology shows that improvements in jurisdictional status can occur with supervisory control in the absence of structural dominance. The example of physiatry shows that jurisdictional superiority can exist with

structural dominance in the absence of exclusive supervision. But in physiatry's case, jurisdictional control was only temporary. When physical therapists moved to expand their occupational domain, the specialty's jurisdictional position quickly collapsed. This course of events makes sense in view of the role professional institutions play in occupational development. It is through its institutions that a profession formulates definitions of jurisdiction and pursues strategies for securing its work domain. Control over occupational structures by another profession would have the effect of imposing an externally generated concept of jurisdiction and of delaying the emergence of indigenous occupational strategies. But in the event that professionalizing movements do arise within the subordinated occupation, structural dominance alone would provide no means for enforcing externally imposed boundaries. To prevent encroachment in the face of an outright challenge, specialists would need leverage to deny jobs to boundary violators. Exclusive supervision provides such leverage by giving the specialty control over its subordinates' labor market.

These observations suggest that structural dominance and exclusive supervision each contribute to jurisdictional control in distinctive ways. Structural dominance appears to secure at least temporary deference to specialists' definition of interprofessional boundaries and to forestall the onset of professionalizing movements among technicians. This may be the optimal route to jurisdictional control during historical periods when supervisory arrangements in the workplace are incompletely institutionalized or undergoing change. Exclusive supervision appears to be a necessary condition for maintaining jurisdictional superiority in the face of boundary-expanding movements by ancillary workers and, for this reason, is crucial to the long-term maintenance of jurisdictional control.

Cognitive and Institutional Imperatives

If receiving a mandate for exclusive supervision was crucial to an emerging specialty's ability to sustain jurisdictional control, what then influenced the willingness of established segments to allocate such privileges? The answer reevokes general themes of the professional dominance and systems perspectives. The historical record points to a number of contributing factors, but two stand out: the subordinating specialty's cognitive legitimacy and the compatibility of its interests with those of dominant medical segments.

In evaluating the emerging specialty's knowledge, professional decision makers were concerned with the credibility of its claim to important expertise not shared by other doctors. The ACS assessments on this matter were consistent with broader professional opinion. Radiology and

pathology had considerable legitimacy within medicine when they received mandates to head hospital departments in the mid-1920s. Both fields had obtained special AMA sections whose creation rested on the vote of the organization's House of Delegates. Reiser (1978) links the medical community's rapid acceptance of radiology and pathology to the perception that the fields provided objective clinical evidence independent of patients' subjective responses. Pathology's clinical wing garnered further status from its basic science branch; physicians considered the latter to be central to medicine's overall knowledge base. The other two segments were in a less enviable position. Anesthesiology and physiatry received approval for AMA sections much later: in 1940 and 1949, respectively. They were also later than radiology and pathology in developing certifying boards and specialty training programs (Stevens 1971). These specialties spent the interwar years struggling with marginal legitimacy.

Physiatry was trying to separate itself from the aggrandized claims made by medical equipment manufacturers during the 1920s (Gritzer and Arluke 1985, pp. 64–65). An ongoing boundary dispute with orthopedic surgery further eroded physical medicine's prospects for favorable supervisory arrangements. Evidence of interspecialty conflict appeared in medical journals beginning in the late 1930s, when physiatrists pushed for exclusive supervision over physical therapists within hospital settings. A partisan for the orthopedists insisted that "physical therapy is more closely associated with orthopedic surgery than any other branch of medicine" (Bagwell 1940, p. 684). The dispute heated up in the 1950s when hospital-based physical therapy departments become more widespread in the aftermath of World War II (Archives of Physical Medicine 1952, p. 299). Leaders of the surgical specialty moved to expand their field's terrain to include "nonoperative orthopedics," an arena that subsumed physiatry's occupational jurisdiction (Gritzer and Arluke 1985, pp. 148–51). A contemporary physiatrist complained: "many orthopedic surgeons consider themselves better qualified to prescribe exercise for their patients than are the physiatrists at their hospitals" (Licht 1956, p. 204). Throughout the middle third of the century, an increasingly influential segment of surgeons straightforwardly rejected physiatry's claims to skills not shared by other doctors.

In the prewar period, anesthesiology also faced problems concerning its professional credibility. Statements made by Karl Meyer, surgeon and medical superintendent of a major urban hospital, illustrate contemporary medical opinion of this specialty's expertise. Speaking at an AMA-sponsored conference in 1935, Meyer (1935, p. 44) argued that the "administration of the ordinary anesthetics . . . is not a highly skilled interpretive science like radiology or pathology, but a relatively noncomplicated, though highly skilled, technical procedure." This assessment

had important implications for supervisory practices. "The active surgeon rarely [could] acquire sufficient knowledge and experience" in the more complex x-ray and laboratory procedures to oversee technicians as competently as the radiologist and clinical pathologist (Meyer 1935, p. 44). But ceding management of anesthesia to a dedicated specialist was unnecessary; the surgeon was fully competent to supervise nurse anesthetists and to handle any anesthesia complications that might arise in the operating room (Meyer 1935, pp. 44–45).

The cognitive legitimacy of anesthesiology rose in the years that followed as surgeons expanded the range of their operative procedures and made increasing use of specialized anesthetic methods—including intravenous, spinal, and regional nerve block techniques. "As anesthesia has become specialized," remarked one surgeon, "so also has there been a growing realization of the need of [medical] specialists in the practice of anesthesia" (Sanders 1945, p. 275; also White 1942). By the late 1930s and early 1940s, many came to view physician anesthesia as essential to the outcome of complex surgical procedures (Maier 1944; Ravdin 1941; Sanders 1945; Stevens 1971, p. 241). These changes were underway when the ACS altered its standardization guidelines to require anesthesiologist department heads.

Cognitive legitimacy was not the only consideration shaping ACS standards on technician oversight. Institutional factors were a second major determinant of supervisory policies. When it came to granting exclusive supervision, it mattered whether the specialty's professional interests were compatible with those of surgery. Radiology and clinical pathology were well positioned in this regard. Among the ACS's goals in initiating hospital standardization was to upgrade surgical practice and to weed out undertrained and incompetent surgeons. The association was especially concerned with limiting the surgery performed by general practitioners. Standardization advocates believed that well-run x-ray and laboratory services would facilitate these aims. Some envisioned hospital-based pathologists and radiologists checking up on diagnoses and monitoring for unnecessary and poorly performed operations (Eisenberg 1925; White 1979, pp. 54–55). Speakers at ACS conferences talked about using laboratories as "a club over the medical men"; there was, after all, "no better source of information as regards the competency of" hospital physicians than that furnished by laboratories (Baldy 1922, p. 10). A physician administrator of the ACS standardization program admonished hospital superintendents to locate laboratory services next to operating rooms so that pathologists would be "at your surgeon's elbow" (MacEachern 1922, p. 25).[7] Others insisted that, when it came to monitoring surgical

[7] Davis (1973, p. 245) discusses MacEachern's work with the ACS hospital standardization program

care, the x-ray plate was "almost equal to an autopsy" and that the radiological lab should also be located "just across the hall from the main operating room" (McReynolds 1925, p. 66).

If institutional imperatives favored exclusive supervision for radiology and pathology, they worked against physiatry and anesthesiology. Physiatry faced a problem of institutional relevancy. The appropriateness of its modalities to the care of long-term and chronic patients left physiatry outside the priorities of acute-care medicine. The specialty's status rose during and after the two world wars when its services were employed in the treatment of injured soldiers. Leaders of the military medical corps promoted physiatry's development, and the specialty held a secure position within the army and the Veterans Administration. But it was unable to transfer this standing to the civilian health care system. Physical medicine remained incompletely institutionalized in community hospitals. Relevancy was not a problem for physiatrists' medical competitors, orthopedic surgeons.

Institutional factors that affected anesthesiology concerned the organization of services. Ceding the management of anesthesia to a dedicated medical specialist implied that a physician anesthetist would be present in the operating room for the duration of surgery. Surgeons had two problems with this arrangement. One concerned reimbursement for services: "The patient would . . . be required to pay an anesthesia fee comparable to the surgical fee" (Meyer 1935, p. 44). Surgeons did not raise the issue of fees when discussing specialist direction for x-ray and laboratory units. Perhaps they assumed radiologists and pathologists would spend less time per procedure, and hence charges would be lower per patient than for physician-directed anesthesia. Whatever the precise sources of the difference, using anesthesiologists was seen as a potential threat to the income of surgeons and hospitals. A department staffed by nurse anesthetists alone was a financial resource for hospitals because anesthesia billings exceeded nurses' salaries. According to anesthesiologists, administrators resisted use of physician specialists in an effort to protect hospital balance sheets (Adriani 1972, p. 666; Bunker 1972, p. 56).

The other problem concerned politics within the surgical suite. Simply put, "there was no place for a second specialist in the power structure of the operating room" (Bunker 1972, p. 144). Surgeons might favor the use of radiologic and laboratory procedures, performed outside the operating room, to monitor the quality and appropriateness of surgery. But ceding authority to another physician during surgery was an altogether different matter. The typical surgeon insisted that he alone "must be in control of the entire surgical team, and the decisions incident to all procedures during the course of the operation must emanate from him"

(Bachmeyer 1935, p. 45). Nurse anesthetists willingly followed surgeons' orders; many were loyal to individual surgeons with whom they had professional relationships of long duration. As one physician anesthetist put it, the surgeon preferred working with nurse anesthetists because the latter were both "cheaper financially and more likely to submit tamely to his whims and eccentricities" (McCormick 1919, p. 67).

Authority in the operating room remained a touchy issue in the 1940s and 1950s, when anesthesiologists began moving into community hospitals in growing numbers. Tyrannical behavior of surgeons was a dominant theme in the professional culture of anesthesiology throughout the middle decades of the century (Bunker 1972; Coe 1970, Lortie 1949). However, by World War II, two other trends were under way that would mitigate institutional resistance to anesthesiologist department heads. First, anesthesiologists were beginning to forge professional roles outside the operating room. With increasing frequency during the war and postwar years, they were overseeing recovery rooms and functioning as consultants to resuscitation, oxygen therapy, and inhalation services (Beecher 1962, pp. 44, 48; Dripps 1962, p. 266; Griffith 1942, p. 555; Searle 1984). At some hospitals, they directed intravenous therapy and blood transfusion units (Griffith 1942, p. 555; Volpitto and Vandam 1982, pp. 95, 139).[8] Second, the development of private medical insurance covering physician services undoubtedly allayed surgeons' concerns about competing bills from anesthesiologists. While such insurance did not become widespread until the 1940s, it was clearly an upcoming trend by the late 1930s (Starr 1982). Thus, when the ACS altered its policy on the supervision of anesthesia services, a combination of cognitive, institutional, and economic factors was eroding surgeons' resistance to anesthesiologist department heads.

CONCLUSION

Historical evidence shows that both cognitive factors emphasized by Abbott and political and institutional factors stressed by power theorists affect intraprofessional decision makers, and through them, cross-professional boundaries. The cases examined here do not allow assessment of the relative importance of intellectual and institutional conditions. Among the four medical specialties included in the analysis, cognitive and institutional imperatives covaried: segments with the more

[8] By the early 1960s, anesthesiologists had ceded blood banking to clinical pathology but they continued to expand their purview outside the operating room by moving into consulting roles in newly created intensive care units (Beecher 1962, p. 44; Volpitto and Vandam 1985, p 168–69, 126–27).

credible knowledge claims also had goals highly compatible with those adjudicating occupational boundaries. This covariation may be more than an accident of history. Like individuals, established professional segments may look especially favorably on bids for legitimacy made by others whose interests are compatible with their own. It is historical actors' assessment of cognitive legitimacy, not objective characteristics of knowledge itself, that allows occupational groups to achieve favorable jurisdictional settlements. A number of interacting factors influence a medical specialty's success at dominating ancillary workers: its cognitive legitimacy, the salience of its techniques to the broader professional community, sources of payment for services, the organization of its work, and the character of settings where work takes place. These factors come to play through their impact on actors who have their own interests in the outcome of jurisdictional struggles.

Specific predictions of the three theoretical traditions receive mixed reviews. The professionalization framework implies that attributes of the subordinated occupations will account for gradients of jurisdictional superiority. But characteristics of the ancillary professions fail to explain patterns of physician control over the four health occupations. Professional power theories suggest that jurisdictional dominance will vary with the extent of physicians' hold over subordinates' occupational structures. This prediction receives considerable support from historical data. Specialties that dominated technicians' professional institutions in the 1920s and 1930s exerted high levels of jurisdiction control by World War II. But structural dominance appears less important than an intraprofessional mandate for favorable supervisory arrangements to the long-term maintenance of jurisdictional control. The systems model predicts that the relative merits of the contenders' intellectual foundations will determine the outcome of jurisdictional struggles. But the actors adjudicating these boundary settlements took little notice of how specialists' knowledge compared with technicians'. At issue was whether specialists' expertise was so consequential and so superior to that of other physicians to warrant interposing specialists between the technician and the referring practitioner.

The central argument of this paper lies outside the three major perspectives in the sociology of professions. Instead, I have argued here that the dynamics among a profession's internal segments are critical in shaping its cross-professional boundaries. An emerging specialty's success at subordinating ancillary workers is contingent upon support from established segments and the broader professional community. Jurisdictional control involves two interdependent levels of professional relations; subordination on one level begins with collaboration at the other. This analysis suggests that a fundamental corrective to Abbott's framework is neces-

sary, at least when it is applied to medical professions. Abbott (1988) mentions specialization as a phenomenon that can strengthen or weaken jurisdiction (pp. 106–7). But his guiding metaphor is that of a chain of professions; what happens in one profession is linked by "chains of effects" to what occurs in other professions (pp. 88–90). The cases explored here show that what happens between intraprofessional segments is also vital to a system of professions. We need a model more complex than that of a chain of professions, one that captures the importance of both levels of professional relations and their interdependence.

The question remains whether the processes just described apply outside the health care sector. American medicine has at least one distinctive characteristic that may heighten the importance of intraprofessional dynamics. It has wielded an unusual degree of influence over the organization of medical work sites. Between World War I and the early 1960s, physicians held sufficient authority to shape the division of labor within hospitals. During the same historical period, private locales prevailed over public locales in adjudicating boundary disputes involving medicine. By 1920, medical leaders attempted to keep many of the profession's jurisdictional conflicts out of courts and state legislatures and, for the next four decades, they were largely successful in doing so.[9] My historical account includes two instances where boundary disputes entered judicial or legislative arenas; even these serve to underscore the importance of private locales and intraprofessional dynamics. The first episode is anesthesiology's effort to have the courts declare nurse anesthesia illegal. The specialty failed because surgeons sided with the ancillary profession. A lack of intraprofessional consensus nullified action in the judicial arena. The second episode is physical therapists' move for state licensure immediately after World War II. In this case, ancillary workers were responding to an opening created by the ACS's decision not to prescribe supervisory arrangements in the workplace. This discussion suggests that, where a highly legitimate profession controls the division of labor in professional organizations, private locales (and intraprofessional dynamics) take precedence over public locales (and interprofessional forces) in shaping boundaries between occupations.

By focusing on medicine as an example of a professionally dominated work site, I have chosen cases that demonstrate the impact of intraprofessional processes in its starkest terms. Abbott remarks that "large, profes-

---

[9] The disputes that leaders were most successful in keeping out of legislative and judicial arenas were those between medical specialties and between specialties and newly emerging ancillary professions. Conflicts between medicine and neighboring occupations of 19th-century origin (e.g., pharmacy, ophthalmology, and nursing) continued to take place, as they had before World War I, in public arenas (Halpern 1991)

sionally dominated work organizations that organize specialized professions into an overall division of labor" are no longer restricted to medicine; they are also emerging in law, accounting, and architecture (1988, p. 107). Patterns identified here thus may be increasingly relevant to other professions.

REFERENCES

Abbott, Andrew 1988. *The System of Professions*. Chicago: University of Chicago Press.

Adriani, John. 1972. "Four Decades of Association with the Pioneers of Anesthesiology." *Anesthesia and Analgesia* 51 (5): 665–70.

Albrecht, Gary, and Judith Levy. 1982. "The Professionalization of Osteopathy. Adaption in the Medical Marketplace." Pp. 161–206 in *Research in the Sociology of Health Care*, vol. 2. Greenwood, Conn . JAI Press.

American College of Surgeons. 1925 "Presentation of the Official Report of Hospital Standardization for the Year 1924." *Bulletin of the American College of Surgeons* 9 (1). 7–16.

———. 1933. *Hospital Standardization Report for the Year 1932*. Chicago. American College of Surgeons.

———. 1938. *Manual of Hospital Standardization* Chicago. American College of Surgeons.

American Medical Association (AMA). 1936. "Medical Education in the United States and Canada." *Journal of the American Medical Association* 107 (9): 661–92.

———. 1942. "Hospital Services in the United States." *Journal of the American Medical Association* 118 (13). 1053–1142.

———. 1947. "Hospital Services in the United States." *Journal of the American Medical Association* 133 (15): 1065–1154.

———. 1948. "Hospital Services in the United States." *Journal of the American Medical Association* 137 (16). 1385–1466.

American Registry of Physical Therapy Technicians. 1950. "1950–1951 Directory." *Archives of Physical Medicine* 31:311–51.

American Society of Radiographers, Council on Education and Registration 1932. "First Annual Report of the Council on Education and Registration." *X-Ray Technician* 4 (1). 8–10.

Archives of Physical Therapy, X-Ray, and Radium. 1926 "Commercialized Roentgenology." *Archives of Physical Therapy, X-Ray, and Radium* 7 493

———. 1927 "The Future of Physical Therapy." *Archives of Physical Therapy, X-Ray, and Radium* 8:86–87.

———. 1934 "Registry of Technicians—a Pressing Problem." *Archives of Physical Therapy, X-Ray, and Radium* 15.241–42.

Archives of Physical Medicine 1952 "Is Physical Medicine Oversold?" *Archives of Physical Medicine* 33.299–301.

Archives of Physical Medicine and Rehabilitation. 1953. "Medical Supervision of Physical Therapy." *Archives of Physical Medicine and Rehabilitation* 34:501–2.

———. 1954. "Direct Medical Supervision of Physical Therapy." *Archives of Physical Medicine and Rehabilitation* 35:99–100

———. 1955. "Teamwork in Medicine." *Archives of Physical Medicine and Rehabilitation* 36:43–45.

Bachmeyer, A. C. 1935. "Should the Radiologist. the Pathologist, and the Anesthesiologist be Licensed to Practice Medicine: The Point of View of the Hospital Administrator." Pp. 45–47 in *Proceedings of the Annual Congress on Medical Education,*

Hospitals, and Licensure, February 18–19. Chicago: AMA Council on Medical Education and Hospitals.

Bagwell, Troy. 1940. "Physical Therapy in Relation to Orthopedics." *Archives of Physical Therapy* 21:684–87.

Baldy, John M. 1922. "The Relation of Hospital Standardization to Laboratory Service." *Bulletin of the American College of Surgeons* 6 (2): 10.

Ballinger, Carter M. 1962. "Medical Anesthesia Is Forging Ahead." *Clinical Anesthesia* 2:126–47.

Bankert, Marianne. 1989 *Watchful Care: A History of America's Nurse Anesthetists.* New York: Continuum.

Beecher, Henry K. 1948. "The Nurse Anesthetist." *Surgery, Gynecology, and Obstetrics* 86:115–18.

———. 1962. "Trends in Anesthesia." *Journal of the American Medical Association* 180 (1): 43–48.

Begun, James, and Ronald Lippincott. 1987. "The Origins and Resolution of Interoccupational Conflict." *Work and Occupations* 14 (3): 368–85.

Bendixen, H. H. 1967. "Anesthesiology Manpower Shortage—Fact or Fiction?" *Clinical Anesthesia* 2.16–21.

Betcher, Albert. 1982. "Historical Development of the American Society of Anesthesiologists, Inc." Pp. 185–211 in *The Genesis of Contemporary American Anesthesiology,* edited by Perry P. Volpitto and Leroy D. Vandam. Springfield, Ill.: Thomas.

Betcher, Albert M., Benjamin J. Ciliberti, Paul M. Wood, and Lewis H. Wright. 1955. "Fifty Years of Organized Anesthesiology." *Journal of the American Medical Association* 159 (8): 766–70.

Brecher, Ruth, and Edward Brecher. 1969. *The Rays: A History of Radiology in the United States and Canada.* Baltimore. Williams & Wilkins.

Brown, Carol. 1973. "The Division of Laborers: Health Professions." *International Journal of Health Services* 3:435–44.

Bucher, Rue, and Anselm Strauss. 1961. "Professions in Process." *American Journal of Sociology* 66:325–34.

Bunker, John P. 1972. *The Anesthesiologist and the Surgeon.* Boston: Little, Brown.

Burdick, Ward. 1924. "Laboratories—an Adequate Laboratory Service in the Modern Hospital." *Bulletin of the American College of Surgeons* 8 (1): 78–81.

Calmes, Selma H. 1984. "American Anesthesia from 1920 to 1950: Was It Women's Work?" Paper delivered at the American Association for the History of Medicine meetings, San Francisco, May 5.

Carr-Saunders, A. M., and P. A. Wilson. 1933. *The Professions.* Oxford: University Press, Clarendon.

Coe, Rodney. 1970. "Processes in the Development of Established Professions." *Journal of Health and Social Behavior* 11:59–67

Cook, Karen, Patricia J. Moris, and Susan Kinne. 1982. "An Exchange Analysis of Emergent Occupational Groups: The Case of New Health Care Providers." Pp. 1–29 in *Research in the Sociology of Health Care,* vol. 2. Greenwood, Conn.: JAI Press.

Davis, Loyal. 1973. *Fellowship of Surgeons: A History of the American College of Surgeons.* Chicago: American College of Surgeons.

Desjardins, Arthur U. 1931. "The Status of Radiology." *Journal of the American Medical Association* 96 (21): 1749–53.

Dripps, Robert J. 1962. "Decisions for a Specialty." *Bulletin of the New York Academy of Medicine* 38 (4): 264–70.

Eisenberg, Arthur A. 1925. "Errors in Surgical Diagnosis, Avoidable and Unavoidable." *Bulletin of the American College of Surgeons* 9 (1): 46–56.

Etzioni, Amitai, ed. 1969. *The Semi-Professions and Their Organization.* New York: Free Press.

Evans, Patricia R. 1978. "Accreditation—Vehicle for Change. *Physical Therapy* 58.151–55, 439–44.

Freeark, Robert J. 1981. "Presidential Address. Current Relations between Surgery and Anesthesiology—a Look at the Other Side." *Surgery* 90 (4): 565–76.

Freidson, Eliot 1970 *Profession of Medicine* New York: Harper & Row

Goff, Joseph H. 1977. "A Survey of Nurse Anesthetists, 1976—a Condensed Report." *Journal of the American Association of Nurse Anesthetists* 45 (2). 143–62.

Goode, William J 1960. "Encroachment, Charlatanism, and the Emerging Profession: Psychology, Sociology, and Medicine." *American Sociological Review* 25 (6) 902–14

———. 1969. "The Theoretical Limits of Professionalization." Pp. 266–313 in *The Semi-Professions and Their Organization,* edited by Amitai Etzioni. New York: Free Press.

Greene, Alfred B. 1954. "Quo Fata Vocant?" *X-Ray Technician* 26:76–88.

Griffith, Harold R. 1942. "Anesthesia Service for Small Hospitals." *Anesthesiology* 3:553–59.

Gritzer, Glenn, and Arnold Arluke 1985. *The Making of Rehabilitation.* Berkeley and Los Angeles. University of California Press

Gunn, Ira. 1983 "Professional Territoriality and the Anesthesia Experience." Pp 155–68 in *Nursing Issues for the Eighties,* edited by Bonnie Bullough, Vern Bullough, and Mary C. Soukup New York Springer.

———. 1975. "Nurse Anesthetist-Anesthesiologist Relationships Past, Present, and Implications for the Future." *AANA Journal* 43 (4): 129–39

Halpern, Sydney A. 1988. *American Pediatrics: The Social Dynamics of Professionalism, 1880–1980.* Berkeley and Los Angeles: University of California Press.

———. 1991. "Paternalistic and Competitive Strategies of Interprofessional Dominance " Paper presented at the American Sociological Association meetings, Cincinnati, August 26.

Hazenhyer, Ida M. 1946. "A History of the American Physiotherapy Association." *Physiotherapy Review* 26.3–14, 66–74, 122–29, 174–84

Hibben, J. Severly. 1930. "Relation of the Attending Physicians to the Physical Therapy Department of the Hospital." *Archives of Physical Therapy, X-Ray, and Radium* 11:464–69.

Hillkowitz, Philip 1924 "The Relation of the Clinical Pathologist to the Hospital " *Journal of Laboratory and Clinical Medicine* 9 (6). 361–65.

Jackson, Robert M. 1984. *The Formation of Craft Labor Markets.* New York: Academic Press.

Kolmer, John A. 1925. "Education as a Cure for Present-Day Evils in Clinical Pathology." *Journal of Laboratory and Clinical Medicine* 10:891–97.

Kovacs, Richard. 1934. "Place of Physical Therapy in Organization of a General Hospital." *Archives of Physical Therapy, X-Ray, and Radium* 15 685–86

Kronus, Carol. 1976 "The Evolution of Occupational Power." *Sociology of Work and Occupations* 3 (1): 3–37.

LaField, William A. 1923. "Radiological Service." *Bulletin of the American College of Surgeons* 7 (2): 26–27.

Licht, Sidney. 1956. "The Future of Physical Medicine. Specialty Respectability." *American Journal of Physical Medicine* 35:201–5.

Lortie, Dan C. 1949. "Doctors without Patients: The Anesthesiologist—a New Medical Specialty." Masters' Thesis. University of Chicago, Department of Sociology

MacEachern, Malcolm T. 1922. "How Hospital Standardization Can Improve the Professional Work and the Service to the Patient in the Hospital." *Bulletin of the American College of Surgeons* 6 (2). 19–32.

Magath, Thomas B. 1939. "The Practice of Clinical Pathology in Relation to Medical Economics " *American Journal of Clinical Pathology* 9 381–96.

Maier, Herbert C 1944. "Responsibility of the Anesthetist in Reducing the Operative Complications of Thoracic Surgery." *Anesthesiology* 5·11–21.

McCormick, A. S 1919. "Some Observations on Anesthesia as a Specialty and the Anesthetist as a Specialist " *American Journal of Surgery, Anesthesia Supplement* 33 (7): 66–68.

McReynolds, George S. 1925. "The Relation of the Hospital to the Departments of Radiology and Pathology." *Bulletin of the American College of Surgeons* 9 (1). 64–66

Meyer, Karl A 1935. "Should the Radiologist, the Pathologist, and the Anesthesiologist Be Licensed to Practice Medicine The Point of View of the Surgeon " Pp. 43–45 in *Proceedings of the Annual Congress on Medical Education, Hospitals, and Licensure,* February 18–19. Chicago AMA Council on Medical Education and Hospitals.

Miller, Albert H. 1933. "Organization of the Anesthesia Service of the General Hospital " *Journal of the American Medical Association* 101 (15): 1119–21.

Montgomery, Lall G. 1970. "A Short History of the Registry of Medical Technologists of the American Society of Clinical Pathologists." *American Journal of Clinical Pathology* 53:433–46

Murphy, John T. 1933. "Influences Affecting the Future of Roentgenology." *American Journal of Roentgenology and Radium Therapy* 30 718–22.

Pancoast, Henry K 1933. "The Future of Radiology as a Medical Specialty." *American Journal of Roentgenology and Radium Therapy* 30 (6). 711–17

Perry, Gentz. 1930. "Ethics for Radiographers " *X-Ray Technician* 2 (2). 35–39.

———. 1933. "The Proper Education and Registration of X-Ray Technicians." *X-Ray Technician* 5 (2) 75–79.

Ravdin, I. S. 1941. "A Surgeon Comments on the Specialty of Anesthesiology." *Anesthesiology* 2:207–8.

Reiser, Stanley J. 1978 *Medicine and the Reign of Technology.* Cambridge: Cambridge University Press.

Ritzer, George. 1977. *Working: Conflict and Change,* 2d ed. Englewood Cliffs, N.J.. Prentice-Hall.

Robinson, Barbara O. 1956. "The Physical Therapy Profession and the Political Process." *Physical Therapy Review* 36 (9): 391–98.

Rose, Donald L. 1959. "The Practice of Physical Medicine and Rehabilitation." *Archives of Physical Medicine and Rehabilitation* 40 1–7.

Rosen, George. 1944. *The Specialization of Medicine.* New York. Froben.

Rosenberg, Charles E. 1979. "Inward Vision and Outward Glance: The Shaping of the American Hospital, 1880–1914." *Bulletin of the History of Medicine* 53 (3): 346–91

———. 1987. *The Care of Strangers* New York. Basic Books

Rothstein, William G. 1972. *American Physicians in the Nineteenth Century.* Baltimore. Johns Hopkins University Press.

Rothstein, William G. 1979. "Pathology: The Evolution of a Specialty in American Medicine." *Medical Care* 27 (10). 975–88.

Sanders, R. L. 1945. "Surgeon-Anesthetist Relations." *Anesthesiology* 6:273–78.

Searle, Maureen. 1984. "The Professionalization of Anesthesiology—Again." *Social Science and Medicine* 18 (4): 323–27.

Simpson, Richard, and Ida Simpson. 1969. "Women and Bureaucracy in the Semi-Professions." Pp. 196–265 in *The Semi-Professions and Their Organization,* edited by Amitai Etzioni. New York: Free Press.

Simpson, Walter M. 1933. "The Training of Technicians for the Clinical Laboratory." Pp 13–14 in *Proceedings of the Annual Congress on Medical Education and Licensure,* February 13–14. Chicago: AMA Council on Medical Education and Hospitals.

Smith, Marion G. 1936. "The American Registry of Physical Therapy Technicians." *Archives of Physical Therapy, X-Ray, and Radium* 17:619–22.

Starr, Paul. 1982. *The Social Transformation of American Medicine*. New York: Basic Books.

Stevens, Rosemary. 1971. *American Medicine and the Public Interest* New Haven, Conn.: Yale University Press.

———. 1989. *In Sickness and in Wealth*. New York: Basic Books.

Thatcher, Virginia S 1953. *History of Anesthesia with Emphasis on the Nurse Specialist*. Philadelphia: J B. Lippincott.

Volpitto, Perry, and Leroy D. Vandam, eds. 1982. *The Genesis of Contemporary American Anesthesiology*. Springfield, Ill.: Thomas

White, Charles S. 1942. "A Surgeon Looks at Anesthesia." *Anesthesiology* 3·455–56.

White, William D. 1979. *Public Health and Private Gain*. Chicago: Maaroufa.

Uncertainty and Professional Work: Perceptions of Physicians in Clinical Practice[1]

Martha S. Gerrity, Jo Anne L. Earp, and Robert F. DeVellis
University of North Carolina at Chapel Hill

Donald W. Light
University of Medicine and Dentistry of New Jersey

Despite growing awareness of uncertainty in technical and scientific fields, uncertainty among physicians, except among physicians in training, is not well researched. Existing studies have primarily used small samples and qualitative methods. This article reports the first rigorously developed measures of uncertainty administered to a large sample of practicing physicians. In contrast to denial and the tendency to minimize uncertainty reported in field studies of trainees, physicians in this study readily acknowledged uncertainty in a number of areas.

INTRODUCTION

The rapid growth of technology and scientific knowledge has brought with it, paradoxically, a rapid increase in uncertainty. Uncertainty borders the edges of knowledge, so that the larger the territory known, the more extensive are the settings in which uncertainty is experienced. As science and technology push toward ever more fundamental frontiers in molecular biology and thermonuclear physics, our awareness of uncertainty deepens (Sorenson 1974; Douglas and Wildavsky 1982, pp. 49–66; Wolf, Gruppen, and Billi 1985; Slovic 1986). Rationality, as a cultural icon, accentuates the uncertainties at the edges of our newest, most fun-

damental knowledge (Fox 1989, pp. 183–202). Each of its positive attributes—logic, objectivity, value neutrality, universality, reductionism, and determinism—makes ever more problematic those phenomena characterized by happenstance, bias, complexity, contingency, and intuition. As rising expectations and the dramatic accomplishments of science and technology intensify the quest for certainty, ironically our awareness of risk and uncertainty increases proportionately (Bosk 1986; Peterson and Pitz 1988; Kassirer 1989). In clinical settings, too, expectations have grown as the number of serious diseases cured has increased and knowledge expanded; this results in a more painful uncertainty when medicine's limits and doubts are confronted.

Despite a constant stream of articles in the scientific literature about the statistical, philosophical, and cognitive aspects of uncertainty, sociology has barely explored this shadow world of modern science and professional life. Intrinsic to the technical parameters and cognitive experiences of uncertainty are its sociological and cultural contexts. Yet, recent attempts to understand uncertainty come mainly from disciplines that rely heavily on psychological or cognitive frameworks (e.g., Hershey and Baron 1987; Curley, Young, and Yates 1989) and focus on risk and decision making (Hogarth 1987; Hershey and Baron 1987; Slovic 1986, 1987). These perspectives, with their highly rational models such as expected utility theory and Bayesian probability theory, describe or predict but a portion of actual behavior (Wolf et al. 1985) because they ignore social, cultural, and organizational forces (Douglas and Wildavsky 1982, pp. 79–80). The sociology of uncertainty, with its attention to norms, beliefs, rituals, and institutional responses, has yet to be written.

REVIEW OF THE LITERATURE

Foundation Work

Uncertainty about causes and effects in medicine and life have preoccupied many students of society, but perhaps none has matched E. E. Evans-Pritchard's (1937) study, *Witchcraft, Oracles and Magic among the Azande*. Zande witchcraft provides a precise system for addressing the uncertainties of life. For example, if someone falls ill, the Azande can explain not only *how* but *why* that *particular person* was afflicted at a particular moment: someone bewitched him or her (1937, pp. 507–8). There are practically no illnesses of unknown diagnosis, no deaths from unknown causes.

The dual cause of disease among the Azande calls for two lines of attack, one to counter the disease via medicines and leech-craft and one to counter the witchcraft via witch doctors, oracles, and magic (Evans-

Pritchard 1937, p. 541). Evans-Pritchard identified thousands of medicines of extraordinary range, medicines for causing rain to fall, for ensuring good harvests, for hexing witches, and for assuring success in love affairs. The medicinal powers of this enormous pharmacopia are infused through ritual and magic rather than residing in the substance itself. Of course, they did not always work, and Evans-Pritchard identifies 22 classes of explanations for these failures, many of which have to do with timing, the effects of social power, and the countervailing powers of witchcraft.

Evans-Pritchard's study is an exemplar of mapping the responses to uncertainty in the social structure and group dynamics of the community. All illnesses and deaths emanate from and reflect back on the structural contexts and group dynamics surrounding the victim so that, in the end, a limited but rather precisely defined amount of uncertainty remains.

Malinowski (1948), in an influential work, also addresses uncertainty, although he does so implicitly by explaining how magic, science, and religion differ as ways to know and affect the unknown. Primitive man experiences impasses "where gaps in his knowledge and limitations of his early power of observation and reason betray him at a crucial moment" (p. 70). According to Malinowski's psychological theory, the natives react to these uncertainties by so intensely obsessing about the desired end through ritual that they imbue themselves with the magical power to achieve it (pp. 60–70). In Malinowski's view, "from our high places of safety in developed civilization" magic is based on emotional states in man and the belief "that hope cannot fail nor desire deceive," while science is "founded on observation, fixed by reason" (pp. 60–70). Malinowski's failure to accept magic systems as complex, integrated constructions of reality on their own terms influenced much subsequent work.

It was Mary Douglas (1966), in a brilliant theoretical essay, who traced to Robertson Smith, and allegedly thence to Durkheim, the origins of these prejudices and separations of magic from religion. Although Douglas wrote specifically about dirt, pollution, and risk, many of her insights apply to uncertainty as well. Uncertainty does not exist in the absolute but only in relation to order. The more differentiated the order, the more uncertainty appears. Conversely, uncertainty defines what is ordered and known. Elaborate rules and rituals are developed to prevent uncertainty, to minimize it, to attribute responsibility for it, and to eliminate it. In particular, we have replaced symbolic spiritualism with scientific materialism and therefore no longer "see" the symbolic role of our ideas about it. We are not even aware of our own rituals, which frame, aid, and ultimately formulate and modify experience.

In sociology, Parsons (1951) seems to have accepted Malinowski's ob-

servations while avoiding his prejudices; thus he was able to present a broader, more balanced, set of observations. Parsons recognized that magical thinking was by no means restricted to preliterate societies but was found in lay accounts, home remedies, patent medicines, and popular superstitions of modern societies. Applying science to illness is the exception among the world's cultures, even in countries where Western medicine and science dominate (Payer 1988, pp. 23–34). It was his student, Renée Fox (e.g., 1976b), who made the final inference, that scientific medicine is itself a symbolic system for coping with the fears and uncertainties of medicine.

Parsons (1951) recognized that scientific medicine ups the ante for doctors because it places on them the full burden of diagnosis, effective treatment, and the control of social disruptions caused by illness. He described at length the burdens and stresses of the doctor-patient relationship and "the element of uncertainty which looms so large in medical practice" (p. 566). Yet "the physician's responsibility is to 'do everything possible' to forward the complete, early and painless recovery of his patients" (p. 450). No Zande ever expected so much from even the best witch doctor or curer.

Perhaps the best explication and extension of Parsons's ideas about uncertainty and trust was made by Bidwell (1976). He systematically delineates classes of uncertainty arising from aspects of professional work: conflicts between the ethos of scientific inquiry and the application of scientific knowledge for practical problems, the indeterminacies of using scientific knowledge for practical purposes, the uncertainty over who controls the application of scientific knowledge arising from communication gaps between laymen and experts, and the tensions between science and other bases of legitimacy. Trust, according to Bidwell (1976), is central to solving these problems and is inherent in specialization itself. Without it, the whole system of complementary roles and tasks that allow professional help collapses.

Studies of Uncertainty in Medicine

Against the rich backdrop of anthropological and sociological thinking, there are few modern empirical studies by sociologists of uncertainty. Most interesting is the work by Renée Fox. Starting with field observations, interviews, and student diaries, Fox (1957) documented the medical school experiences that led students to appreciate three types of uncertainty: that arising from incomplete mastery of the vast and expanding range of medical knowledge and skills, uncertainty due to the limitations and ambiguities of that knowledge and those skills, and the uncertainties of distinguishing between the first two. In addition to describing these

types of uncertainty, Fox noted that stress resulting from a sense of "personal inadequacy" (pp. 210–11, 231–33) was a prominent characteristic of medical students' reactions, both adaptive and maladaptive, to uncertainty.

Fox is the only sociologist to have made uncertainty a central focus of sustained work, and she did so in the special context of research-clinicians (Fox 1959, 1980, 1989; Fox and Swazey 1975). Her research did not yield a systematic analysis of types of uncertainty, but it did produce a full catalog of ways in which both doctors and patients cope in this special setting. Most commonly cited are the game of chance, in which physicians pit their professional skills in friendly competition and at the same time practice magical thinking, and "gallows humor," with its ability to lighten the burdens of ignorance and powerlessness. In *Experiment Perilous,* Fox (1959) identifies a host of other, too-often overlooked, coping mechanisms. They include group bonds and commitment to each other; caring, informing, and sharing with patients; honoring patients for their courage and dedication; telling over and over stories of success as myths of the enterprise; trying out new procedures, drugs, or combinations on a hunch; keeping meticulous records in hopes of discovering patterns in the uncertain swirl of events; turning to religion; and accepting ignorance, uncertainty, and death as part of life.[2] Given Malinowski's perspective on magic, it took unusual intellectual integrity for Fox to see magical thinking at the center of scientific work and to treat it with respect.

Other forms of what Fox characterizes as "scientific magic" include exuberantly focusing on the positive features of an experimental therapy; "the necessity to hope"; and the custom of giving an animal a name a day or two after an implant as if to confirm that it will live (Fox 1976a, 1984; Fox, Swazey, and Cameron 1984). When magic does not "work" and things go badly, researchers announce a "clinical moratorium" as a collective retreat and stocktaking (Swazey and Fox 1970; Fox 1989, pp. 198–202).

Aside from this body of work, most sociological research on uncertainty in medicine has focused on those in training. As research shifted from medical school to residency training, new kinds of uncertainty were identified. Light (1979, pp. 311–12), in consultation with Fox, identified "a new bundle of uncertainties . . . as the young professional takes on actual cases." Cross-cutting the three uncertainties of knowledge described in Fox's studies of medical school are clinical uncertainties surrounding diagnosis, treatment, client responses and supervisors' reactions. Ac-

---

[2] Detached concern is presented but seems to dissolve before the many forms of attached concern (Fox 1959, pp. 86–104).

cording to Light (1979), these exist in specialties as divergent as psychiatry and orthopedic surgery.

Light (1979) emphasized that, while students learn to appreciate the uncertainties of knowledge in what are called the "undergraduate years" of medical school, as residents they must learn how to control the many sources of uncertainty or be paralyzed by it. Hence, residents learn forms of control matched to each kind of uncertainty, and each form of control builds on those before it. To control the uncertainties of subordination, residents "psych out" supervisors and conform to their expectations (also see Bosk 1979, 1980). To control uncertainties of knowledge, residents reduce the problem by specializing and adopting a "school of thought" that resolves uncertainties through conviction; and in so doing, they master a great deal of information. To manage uncertainties of diagnosis residents use the above techniques, plus they invoke the clinical experiences they acquire as a ritual anecdotal resolution. Uncertainties of treatment are controlled by all the preceding strategies but with an emphasis on technique that transposes adverse outcomes into correct procedures or treatments. Finally, the uncertainties of patient response are reduced by the above techniques *and* by limiting what the patient knows (see also Waitzkin and Waterman 1974). Perpetuating and reshaping patient uncertainty helps protect physicians from being confronted with their own uncertainties, and makes their job easier, as Fred Davis (1960) found among doctors treating victims of polio.

Ironically, withholding information results not infrequently in heightening patients' uncertainties possibly since most patients want as much information and feedback as they can get (Waitzkin 1983, 1985). Waitzkin, a practicing physician as well as a sociologist, has conducted extensive research on ways in which physicians, managing their own uncertainty, maintain power and control through keeping patients unaware of diagnostic information or therapeutic options. On the other hand, research addressing the effects of physician uncertainty on patient satisfaction (Johnson et al. 1988; Gutheil, Bursztajn, and Brodsky 1984) and treatment decisions (Curley, Eraker, and Yates 1984) suggests that patients may be less satisfied with their care when physicians express their uncertainty, and in some cases patients prefer to defer ambiguous treatment decisions to physicians.

Bosk (1980), in a major study of surgical residency, found that residents used eight rituals to manage the uncertainty of diagnostic and treatment decisions: hedging assertions, probabilistic reasoning, requests for consultations, Socratic teaching, deciding not to decide, gallows humor, hyperrealism, and focusing on uncertainty as a research problem. And Atkinson (1984) argues that physicians learn to manage uncertainty by receiving a "training for certainty." They learn to view the science

underlying medicine as "established 'facts' and soluble 'puzzles.'" From this perspective, patients' problems always result from identifiable diseases, and, therefore, once the "correct" identification has been made, treatment and related recommendations automatically and predictably follow.

Light (1979) expressed concern that this "training for control" (or Atkinson's "training for certainty"), when successful, leads to premature closure and denial of genuine uncertainties in clinical work. At the center of this training is the top-down tradition of education based on a class hierarchy that judges students and trainees by their ability to "know the answers" and do things "right" (Bowers 1987, pp. 49–72). What may be covered over is the scientific uncertainty that actually exists. As Bosk (1979) discovered, normative errors arising from failure to conform to "the way things are done" by a resident's attending physician are far more serious than technical errors because they indicate professional insubordination and possible irresponsibility.[3] Light (1979) points out that a "school of thought" is an ideological resolution of uncertainty. This is aptly illustrated by a physician who trained under a renowned advocate of radical breast surgery and exclaimed, "Anything but radical mastectomy is criminal conduct!" when, in fact, the evidence is equivocal and complex (Katz 1984, p. 39).

Strong defenses against and denial of uncertainties is one of the most consistent observations made by sociologists studying medical training. Bucher and Stelling (1977), in their comparative study of residents in several specialties, concluded that by the end of the second year residents could no longer "hear" or learn from alleged mistakes. Light (1980, chap. 4, pp. 297–307) came to a similar conclusion in his discussion of the ways in which the structure of residency encourages feelings of omnipotence. More recently, Mizrahi (1984, 1986) has emphasized that residents are given no training in how to cope constructively with uncertainty and, in response, they acquire elaborate rituals for coping with mistakes or defending themselves against uncertainties. She found that medical faculty have become less involved in day-to-day patient care, hence residents who are without role models invoke even stronger defenses when faced with uncertainty.

Katz (1984, p. 37), in his influential writings, emphasizes "how pervasive the disregard of uncertainty becomes whenever uncertainty ceases to be merely theoretical and impinges on . . . actual clinical encounters." In fact, Katz concludes that, once in practice, physicians are reluctant

---

[3] For a systematic outline of adaptations to uncertainty, see Vyner (1988), Curley, Young, and Yates (1989), and Curley et al. 1984.

to admit Fox's three kinds of uncertainties of knowledge that they previously recognized and appreciated during their training. Indeed, the denial of uncertainty allows physicians to make potentially threatening situations more understandable and controllable, thus enabling action to take place.

Denial of uncertainty, defenses against criticism, and management of patients to maintain dominance, if they are common products of training, have important consequences for medical practice. They may be predisposing factors behind overtreatment and iatrogenesis, diseases and health problems caused by medical treatment (Applegate 1986). "The high rate of 'unnecessary' surgery, of resort to antibiotics and to tranquilizers, bears testimony to physicians' propensity to resolve uncertainty and ambiguity by action rather than inaction" (Katz 1984, p. 41). This seems equally true of overtesting, which strives for diagnostic certainty rather than pragmatic decisions about treatment (Kassirer 1989; Hardison 1979; Sox 1986). At a time of great concern about containment of health care costs, uncertainty is thought to be an important cause of excessive use of resources (Detsky et al. 1981; Wennberg, Barnes, and Zubkoff 1982) and variations in physicians' practice patterns (Eddy 1984).

It is important to remember that such behavior does not evidence disregard for patients or the acceleration of medical costs but shows instead an intense preoccupation with uncertainty and an insidious aspect of the powerful influence that norms, beliefs, and institutional responses requiring greater certainty hold over physician's behavior. Current efforts to limit the rapid increase in medical expenses only exacerbate the dilemma. The pressure for diagnostic certainty and the push for therapeutic results at any price play themselves out in a world of limited resources but seemingly unlimited technological discoveries (Eddy and Billing 1988).

Although uncertainty is frequently mentioned in the medical literature (Lusted 1984) and past observations and insights have provided an initial framework for approaching an empirical study of uncertainty, the actual patterns of uncertainty among practicing physicians remain unknown (Katz 1984). Previous research focused on physicians in training and used primarily qualitative research methods. This article reports one of the first attempts to move from fieldwork and theory to using quantitative methods to explore, in a large sample of community-based practitioners, physicians' reactions to uncertainty in patient care. The study had three objectives: (1) to begin the process of defining the construct, affective reactions to uncertainty, and developing quantitative measures of the main dimensions of the construct, (2) to develop a conceptual model of factors influencing physicians' reactions to uncertainty as a heuristic device to guide our research in this area, and (3) to map the extent and

variation of physicians' affective reactions to the uncertainty inherent in clinical work.

METHODS

Conceptual Model

Semi-structured interviews with physicians, nurses, and laboratory and insurance personnel and a review of the medical sociology, psychology, and health-services literatures served as a basis for developing a model to identify factors affecting physicians' reactions to uncertainty and how reactions to uncertainty might influence their behavior (fig. 1). The model focuses on five major elements: the patient, the medical problem or illness, the physician, test and treatment characteristics, and the organizational structure. It attempts to integrate internal processes, cognitive and affective, with external influences, sociological and cultural.

In the model, patients present to physicians with medical problems. Besides the medical problem, other characteristics of patients (their age,
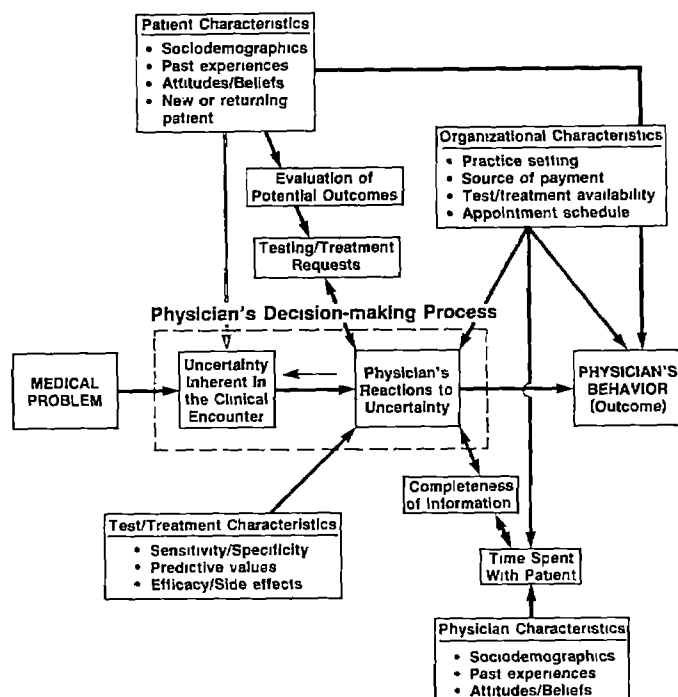

FIG. 1.—Conceptual model of factors influencing physicians' reactions to uncertainty.

gender, socioeconomic status, attitudes toward medical care, need for certainty, social support) may influence physicians' decision making (Eisenberg 1979). The medical problem, together with patients' characteristics, creates the uncertainty inherent in the clinical encounter. Physicians bring to clinical encounters their own reactions to clinical uncertainty, both cognitive and affective. Physicians' reactions may vary depending on their characteristics (age, specialty training, prior patient-care experiences, attitudes and beliefs). Patients and physicians, each with their own set of characteristics, interact to produce a series of medical decisions (test ordering, referral, treatment, follow-up visits, reassurance, hospitalization). The decision outcome, the physician's behavior, may be modified either by system characteristics (practice setting and payment plan), or patient characteristics, or both (patients refuse treatment, insurance companies or prepayment plans deny payment for treatments or hospitalizations, the diagnostic tests are not available).

Definition of the Construct

After developing the conceptual model, we defined the construct "affective reactions to uncertainty in patient care." We discovered from our interviews and literature review that physicians have a variety of reactions to uncertainty, including tolerance of uncertainty. These reactions fell into the following nine categories: patient-physician relationships, physician-colleague relationships, professional norms, self-esteem as a physician, bad outcomes, missed diagnoses, malpractice worries, patient referrals, and test ordering. Of note is the fact that many of the categories involve perceived consequences of uncertainty and possible coping behaviors. We then broadly defined affective reactions to uncertainty in patient care as (a) the emotional reactions and concerns engendered in physicians who face clinical situations that are unfamiliar or not easily resolved and (b) the behaviors used by physicians to cope with those emotions and concerns.

Development of the Instrument

The development of the original pool of 61 items and the final scale has been described in detail previously (Gerrity, DeVellis, and Earp 1990). The methods described here include the physician sample, a summary of the scale development, and analyses pertinent to the data presented in this article.

   The physician sample.—We administered a 61-item scale to a sample of physicians similar to those who, in the future, might be studied using this scale. For this initial sample, we wanted community physicians who

did primary and consultative patient care, had exposure to a wide variety of patient problems, and represented several different specialties and subspecialties. We selected a random sample of 700 physicians, stratified by specialty, from two sampling frames; half were licensed in North Carolina and half in Oregon. Four hundred were generalists in family medicine, general practice, or general internal medicine, 200 were internal medicine subspecialists, and 100 were surgeons.

Of the 700 physicians contacted, 428 returned completed questionnaires, a response rate of 61%. Eleven percent of the nonresponses were due to (1) the physician's being retired for over two years (4%) or out of the country (1%) and (2) no forwarding address (6%). Further attempts to obtain responses from the remaining 241 nonrespondents were not made owing to the length of the questionnaire and the expense and time required by such an extensive follow-up.

The stratified selection of physicians produced a sample whose mean age was 46 years (SD 13 years), 374 (88%) of whom were male, and for whom a mean of 20 years (SD 14 years) had passed since medical school graduation. By design, the majority of the physicians were generalists: 110 (26%) family physicians, 30 (7%) general practitioners, and 66 (15%) general internists. One hundred twenty-three (29%) were internal medicine subspecialists and 43 (10%) were surgeons. The 53 remaining physicians included 3 in residency training, 26 no longer in practice, and 22 who retrained in another specialty such as rehabilitation medicine, radiation oncology, or occupational medicine.

Physicians in our sample considered themselves primary care providers. Three hundred thirty-five (78%) were in full-time practice and 205 (54%) spent 75%–100% of their time in primary patient care. Two hundred seventy-three (73%) of the physicians were in private-group or solo practices and 220 (61%) had 2–15 physicians in their group.

Exploratory factor analysis and selection of items for the final scale.—We asked these physicians to respond to the 61 items using a six-point Likert scale ranging from strongly agreeing with the statement ("1") to strongly disagreeing ("6"). The sample also rated their perception of the amount of uncertainty encountered by physicians in 14 specialties and subspecialties using a 10-point Likert scale ranging from the least amount of uncertainty ("1") to the greatest amount of uncertainty ("10").

We then factor-analyzed the physicians' responses to the 61 items. Exploratory factor analysis allowed us to determine how many dimensions (factors) underlay the entire set of uncertainty items and to identify which items were strongly and unambiguously associated with these factors. (Nunnally 1978, pp. 327–436) A scree test (Cattell 1966) applied to

the factor-analytic results indicated that two factors were adequate to account for the observed findings.

The two-factor solution in our study explained 58% of the common variance in the physicians' responses to the 61 items. Items were evaluated as having a strong and unambiguous relationship with only one factor if their loading on that factor was greater than .40 while their loading on the other was under .20. Items meeting these criteria defined two reliable and readily interpretable scales, which we call the Physicians' Reactions to Uncertainty (PRU) scales. One of these, the final Stress from Uncertainty scale, contains 13 of the original 61 items (mean score = 44, SD 11, α = .90). The other, the Reluctance to Disclose Uncertainty to Others scale, contains nine of the original 61 items (mean score = 23, SD 6, α = .75).[4]

Analysis of the 61 Items and Scale Scores

To assess the extent that physicians in the sample varied in their responses to the original 61 items or held uniform views, we dichotomized their responses into agreeing or disagreeing with each item.[5] We then defined a "consistent response" among physicians as items where 75%–100% of the physicians had the same response (either agreeing or disagreeing with the item). Items that had a consistent response from the study physicians were examined for patterns in their content (patient-physician relationship, malpractice worries, test ordering).

Next, we used the two Physicians' Reactions to Uncertainty (PRU) scales, developed from the 61 items, to explore past assumptions about physicians' reactions to uncertainty and to identify differences among subgroups of physicians along the dimensions Stress from Uncertainty and Reluctance to Disclose Uncertainty to Others. These scales provided us with measures that were more reliable than any single item would be. Because of past observations by Fox (1957) and Budner (1962) and research on gender differences described by others (Bickel 1988; Spiegel, Smolen, and Jonas 1986; Woodward and Adams 1985; Young 1987; Bowman and Allen 1990), we focused primarily on the relationship between physicians' gender, years in practice, specialty group, and their scores on the two PRU scales. Specifically, we examined the hypotheses that:

---

[4] For the Stress from Uncertainty scale, the higher the score the greater the stress. For the Reluctance to Disclose Uncertainty to Others scale, the higher the score the greater the reluctance to disclose uncertainty.

[5] Agreement = strongly, moderately, and slightly agreeing. Disagreement = strongly, moderately, slightly disagreeing with an item based on a six-point Likert scale

(1) physicians' reactions to uncertainty differ as a result of their gender and specialty group and (2) the longer physicians are in practice the less they will react to uncertainty. The *t*-test and analysis of variance (ANOVA) were used to compare mean scale scores (SAS Institute 1985). When ANOVA results indicated a significant difference among the mean scores, individual pairs were compared using Tukey's test (Keppel 1982; SAS Institute 1985). Multiple linear regression, with dummy coding for categorical variables, was used to control for potential confounding among gender, years in practice, and specialty while assessing each variable's relationship to the PRU scale scores (Kleinbaum, Kupper, and Muller 1988, pp. 102–43, 260–96).

Last, physicians' ratings of the amount of uncertainty encountered in the "day-to-day work" of 14 different specialists and subspecialists were examined by first using descriptive statistics. Then, we dichotomized physicians' responses to the 10-point Likert scale (1 = least amount of uncertainty; 10 = greatest amount of uncertainty) into "less uncertainty" or "greater uncertainty" encountered based on the median rating for that specialty. We divided physicians who rated the level of uncertainty encountered in their own specialty (seven of the 14 specialties listed) into two groups based on the median rating of the amount of uncertainty for their specialty. We then compared the mean Stress from Uncertainty scale scores and Reluctance to Disclose Uncertainty scale scores for the "less uncertainty" and the "greater uncertainty" groups using the *t*-test.

FINDINGS

Affective Reactions to Uncertainty

The two dimensions that emerged from the exploratory factor analysis are consistent with our prediction that the construct "affective reactions to uncertainty" would have both an emotional and behavioral dimension. Other components, aside from Stress from Uncertainty (an emotional reaction) and the Reluctance to Disclose Uncertainty to Others (a coping behavior), may exist; indeed "test ordering" was suggested by the three-factor solution in the exploratory factor analysis. However, the breadth of the domain we defined and the number of items we used in our original item pool make it unlikely that our analysis missed major dimensions of the construct. Moreover, other concepts or components, such as malpractice worries and test-ordering, fit within the construct we label "affective reaction to uncertainty," since these also embrace emotional reactions or coping behaviors, the main dimensions of affective reactions to uncertainty. By giving form to this highly abstract construct, this conceptualization should prove useful in guiding future research.

Consistency Among Physicians' Responses

Although the quantitative technique, exploratory factor analysis, defined main dimensions, examining physicians' responses to the 61 items in the original item pool and the extent that physicians had consistent responses to single items gives insight into the variability of physicians' responses to uncertainty and the factors that may modulate these responses. However, caution must be used when examining single items since subtle changes in wording can evoke different patterns of responses (Tversky and Kahneman 1981; McNeil et al. 1982). As one physician in the study noted, "Very interesting. Placing the same question in [a] different format draws a different answer."

The most striking feature of these data is the extensive variability in physicians' reactions to uncertainty. This is evident in both the scale scores, with rather large standard deviations around the mean scores (Stress from Uncertainty scale, 44, SD 11; Reluctance to Disclose Uncertainty to Others scale, 23, SD 6), and the responses to individual items. When physicians' responses to the 61 items were dichotomized into "agree" and "disagree," 75%–100% of the physicians agreed with only 12 items and disagreed with only 11 items (we call these "consistent responses" and list them in App. A, pt. 1 below). For the majority of items (38 of the 61), there was variability in physicians' responses with some agreeing and some disagreeing (these "inconsistent responses" are given in App. A, pt. 2 below).

An example of this variability, án area for research, are the items involving malpractice. With growing concern over malpractice, many physicians feel honesty is the best policy. More than 75% of physicians disagreed with statements such as "I am afraid other physicians would doubt my ability if they knew about my patient care mistakes" and "If I share my uncertainties with patients, I will increase the likelihood that I will be sued." On the other hand, there was greater variability in physicians' responses (25%–74% who agreed) to statements such as "I worry about malpractice when I do not know a patient's diagnosis" and "I frequently think about the legal implications of missing a diagnosis" and "I fear being held accountable for the limits of my knowledge." In addition, more than 75% of the study physicians agreed that, "When physicians are uncertain of a diagnosis, they should share this information with their patients."

Yet, fewer physicians (25%–74%) agreed about what they *actually* do: "I always share my uncertainty with patients." One physician specifically noted that doctors cope with uncertainty individually and in their own way. "Certainly this is a very important subject. Physicians, particularly those out of training, don't share very much with each other the

stress and uncertainty they feel. We don't discuss it and we deal with it each in our own way."

The gap between believing in the honest sharing of uncertainty and actually doing it needs to be measured by more field research. In addition, further research is needed to better define our social concepts of normality since "blameworthiness takes over at the point where the line of normality is drawn" (Douglas and Wildavsky 1982, p. 35).

Explaining Variable Reactions to Uncertainty

To understand the reasons why physicians vary in their affective reactions to uncertainty, we used the two PRU scales, Stress from Uncertainty and Reluctance to Disclose Uncertainty to Others, to compare subgroups of physicians based on certain demographic, professional, and practice characteristics.

Physicians' scores on the Stress from Uncertainty scale distinguished physicians based on their gender, years of experience, and specialty group (table 1). Practicing male physicians reported significantly less stress from uncertainty than did female physicians as evidenced by their lower mean score, 43 (SD 11) versus 50 (SD 10). Additionally, physicians who had been in practice longer felt less stress from uncertainty than their colleagues who had practiced for a shorter time. Specifically, physicians in practice for 17–31 years, as well as those in practice over 31 years, both had significantly less stress produced by uncertainty than physicians in practice less than nine years, based on Tukey's test of mean scale scores. Although Stress from Uncertainty did not differ between internal medicine subspecialists and surgeons, both groups reported significantly less stress from this source than did generalists. Of note is the variability in the scale scores even *within* subgroups of physicians, as evidenced by the relatively large standard deviations.

We made the same comparisons for the Reluctance to Disclose Uncertainty to Others scale. On none of the comparisons did the physicians' mean scores differ significantly. The variation in physicians' scores may be based on other characteristics, such as experience with malpractice and bad outcomes that resulted from uncertainty. These were not examined in our study.

Since female physicians in this study tended to be generalists and have fewer years in practice compared with male physicians, we examined the possibility that gender, years in practice, and specialty were not independently associated with the scale scores but were confounded with each other. Gender, specialty, and region of the country were entered into a linear regression model using dummy coding and years in practice was entered as a continuous variable. Gender ($F = 5.69$, $P < .02$), years in

TABLE 1

COMPARISON OF THE STRESS FROM UNCERTAINTY SCALE SCORES ACROSS GENDER, SPECIALTY GROUP, AND YEARS IN PRACTICE

	N	Mean[a]	SD	Test Statistic
Gender				$t = 4.05$**
Male	373	43	11	
Female	50	50	10	
Years in practice:				$F = 4.37$*
2–9	109	47	10	
10–16	117	44	11	
17–31	97	43	11	
>31	97	42	13	
Specialty groups				$F = 6.21$*
Generalists[b]	213	45	11	
Internal medicine subspecialists	126	42	12	
Surgeons	43	41	11	

[a] The greater the score the greater the stress from uncertainty
[b] Includes general practitioners, family practitioners and general internists.
* $P < 005$
** $P < 0001$

practice ($F = 5.12$, $P < .03$), and specialty (subspecialists vs. generalists and surgeons, $F = 9.07$, $P < .003$; surgeons vs. generalists and subspecialists, $F = 3.76$, $P \leq .06$) each predicted physicians' scores on the Stress from Uncertainty scale after controlling for the other variables in the model. For the Reluctance to Disclose Uncertainty to Others scale, specialty (subspecialists vs. generalists and surgeons; $F = 4.76$, $P < .03$) predicted physicians' scores after controlling for the other variables.

Detailed analyses of the scale scores for other groups of physicians based on characteristics such as practice organization (fee-for-service vs. salary), number of physicians in the practice, or population of the town or city where the practice was located could not be done due to confounding plus small numbers of physicians in each subgroup. For example, men were more likely than women to be in fee-for-service practices (76% vs. 48%); have solo practices (29% vs. 16%); and reside in towns of less than 100,000 people (59% vs. 40%). Physicians in small towns (less than 100,000) were more likely to be generalists or surgeons (78% vs. 52% in larger towns); solo practitioners (37% vs. 14% in larger towns); and in fee-for-service practices (80% vs. 64% in larger towns).

Perceptions of Uncertainty Among Various Specialists

Table 2 describes physicians' responses when asked to rate how much uncertainty they thought physicians in 14 different specialties or subspe-

TABLE 2

Physicians' Perception of the Amount of Uncertainty Encountered in
Day-to-Day Work in 14 Specialties

	Mean Rating of Relative Uncertainty*			
Specialty or Subspecialty	All Physicians (N = 410)	Generalists (N = 205)	Subspecialists (N = 123)	Surgeons (N = 42)
Psychiatry	7 4 (2.2)	7.3 (2 1)	7.6 (2.1)	7 2 (2.5)
Family medicine	6.8 (2.0)	7.0 (2.1)	6 9 (1.9)	6 3 (1 6)
General internal medicine	6.7 (1.9)	6 7 (1.9)	6.8 (1 8)	5.9 (1 8)
Pediatrics	6.0 (2 1)	6.0 (2 0)	6 1 (2.0)	5 4 (2.0)
Obstetrics-gynecology	5.7 (2 2)	6 1 (2 2)	5 2 (2 1)	5.5 (2.4)
Hematology-oncology	5.5 (2.2)	5 5 (2.2)	5.6 (2 2)	4 9 (2 3)
Rheumatology	5.5 (2.2)	5.4 (2.2)	6.0 (2.1)	4 4 (2 1)
Gastroenterology	5.2 (1 8)	5 0 (1 8)	5 5 (1 9)	5.0 (1.9)
Cardiology	5.1 (2.0)	5 0 (2.1)	5.2 (2 0)	5.0 (2.0)
General Surgery	4 9 (1 8)	5.0 (1 8)	4 9 (1.8)	5 0 (2 1)
Dermatology	3.8 (2 2)	3 8 (2.1)	4 1 (2.4)	3 3 (2 0)
Anesthesiology	3 4 (2.1)	3 4 (2.0)	3.4 (2 2)	3.7 (2.2)
Orthopedic Surgery	3.3 (1.8)	3.4 (1 8)	3 2 (1.8)	3 4 (2.0)
Urology	3.2 (1.4)	3 2 (1 4)	3 2 (1.5)	3 2 (1.7)

Note —Numbers in parentheses are SDs
* Ratings were based on a 10-point Likert scale ranging from least amount of uncertainty ("1") to greatest amount of uncertainty ("10")

cialties encountered in their day-to-day work. Overall, physicians felt that psychiatrists encountered the greatest uncertainty followed by physicians in primary care or "generalist" specialties. The surgical subspecialists (urologists and orthopedic surgeons), anesthesiologists, and dermatologists were thought to encounter the least amount of uncertainty in their work. These beliefs persisted even when subgroups of physicians were examined. For example, surgeons felt, as did other physicians, that psychiatrists and generalists encountered greater uncertainty in their work than did surgeons. Surgeons also felt that general surgeons encountered at least as much uncertainty as did internal medicine subspecialists and more uncertainty than surgical subspecialists, anesthesiologists, and dermatologists.

These relative values are not meant to imply that any of the groups of physicians encounter little or no uncertainty in their work or that their impressions of the amount of uncertainty are necessarily correct. As one physician in the study commented, "Before I went into radiation oncology, I thought that it was a field with very little uncertainty—'pretty naive.' I learned that it is filled with uncertainty, as are *all* medical specialties, 'in my opinion.' "

In fact, physicians in one specialty may have little insight into what physicians in other specialties face in their work. This is expressed by another physician in the study. "Basically, I don't feel that physicians in different specialties understand what someone in another specialty really does and the kind of stresses [she or he] deals with."

Indeed, Light's (1979) comparative study of the uncertainty experienced in two widely different specialties (psychiatry vs. orthopedic surgery) found evidence that, from the practitioner's viewpoint, a similar amount of uncertainty was experienced by specialists in both fields.

The perceptions represented by the ratings in table 2 may not be accurate assessments of how much uncertainty physicians in different specialties encounter; however, they do corroborate Fox's (1957) observations of the perceptions that medical students have of the amount of uncertainty and structure in different specialties (Budner 1962). These perceptions of relative uncertainty may be more consequential than the reality; those who perceive more uncertainty in their specialty than in others may report more stress from uncertainty.

To determine how physicians' perceptions of great or little uncertainty in their own specialty related to their reports of perceived stress from uncertainty and their reluctance to disclose uncertainty to others, we divided physicians into two groups based on the median rating of how much uncertainty they felt their specialty encountered in clinical work. Next, we examined the mean Stress from Uncertainty and Reluctance to Disclose Uncertainty to Others scale scores for these two groups. Results for the Stress from Uncertainty scale are shown in table 3. Overall, physicians who perceived that their specialty encountered greater uncertainty reported significantly greater stress from uncertainty than physicians who perceived that their specialty encountered less uncertainty. This trend is present for three specialties, but not present for internal medicine subspecialists, and only the difference in scores for general internists reach statistical significance. These same comparisons were made for the Reluctance to Disclose Uncertainty scale. No significant differences were found among physicians who perceived greater or lesser uncertainty in the practice of their own specialties. The findings on relative perceived uncertainty and stress suggest differences in how uncertainty is handled and experienced within different specialties.

Potential Influences on Reactions to Uncertainty

Physicians appear to have a complex and wide range of reactions to uncertainty. A number of variables, some measured in this study (years of experience, specialty, and gender) and many not measured (severity of an illness, patient's desire for certainty) may shape or influence physi-

TABLE 3

COMPARISON OF MEAN OF STRESS FROM UNCERTAINTY SCALE SCORES FOR PHYSICIANS

Specialty	Less Uncertainty	Greater Uncertainty	Test Statistic
Family physicians	43.5	46 1	$t = 1\ 33^a$
	($n = 32$)	($n = 78$)	
General internists	41.3	48 6	$t = 2\ 83^*$
	($n = 26$)	($n = 47$)	
Internal medicine subspecialists[b]	42.4	41.6	$t = -\ 30^a$
	($n = 35$)	($n = 37$)	
Surgeons	39 3	42.2	$t = .82^a$
	($n = 24$)	($n = 19$)	
Overall group mean	42 2	45.1	$t = 2.21^{**}$
	($n = 117$)	($n = 181$)	

NOTE —The "less uncertainty" and "greater uncertainty" groups were formed based on the median rating physicians gave for the amount of uncertainty encountered in that specialty
[a] N S
[b] Includes rheumatology, cardiology, gastroenterology, and hematology-oncology
* $p < .01$.
** $p < .05$

cians' reactions to uncertainty as described by our conceptual model (fig. 1). Comments made at the end of the questionnaire highlight the complexity of physicians' responses to uncertainty and suggest reasons for the variability between physicians in different specialties, between physicians within the same specialty, and even within individual physicians as they move from one patient encounter to the next.

The medical problem, identified in the conceptual model is one factor that may affect physicians' reactions to uncertainty from one patient encounter to the next. One physician comments on the effect of a dangerous or "high-risk" disease when it is considered in the differential diagnosis (e.g., appendicitis, viral gastroenteritis, etc.) for a medical problem (e.g., abdominal pain): "Many of the questions are highly dependent on [the] situation. For instance, generally [the] diagnosis of appendicitis is not highly stressful because uncertainty is predictable and low risk, especially if [an] operation is undertaken. Other situations are more stressful because the uncertainty leads to much higher risk."

Another example of how the type of medical problem may influence physicians' reactions to uncertainty is seen in the evaluation of patients with sore throats compared to those with chest pain. Missing a diagnosis of a strep throat or treating a patient inappropriately with penicillin will probably not result in a life-threatening situation whereas, with chest pain, there is a risk of missing a heart attack that, if treated inappropriately, could result in death. Uncertainty in the latter situation produces

far greater stress for the physician. Stated another way, by a respondent to our survey: "The emotional response to uncertainty is a function of the consequence of being wrong. I may feel uncertain about the cause of chest pain but comfortable in my uncertainty if important [life-threatening] causes are excluded. There may be many uncertainties in family medicine but my perception is that most of them are not about life-threatening situations."

Consideration of a high-risk disease in the differential diagnosis of a medical problem heightens the stakes and the stress from the diagnostic uncertainty Light (1979) described. The fear or anxiety over bad outcomes resulting from being wrong or uncertain has been described in various ways by Feinstein (1985) and others (Nightingale 1987, 1988; Bell 1982; Hershey and Baron 1987; Williams, Suchman, and Herrera 1988). They have termed these influences the "chagrin factor," "risk aversion," and "regret." In fact, the stress from uncertainty in relation to a medical problem, and its differential diagnosis, may be due to the feelings of personal inadequacy that Fox (1957, pp. 240–41) noted among medical students and to the cultural and institutional frameworks that define "high risk" (Douglas and Wildavsky 1982, pp. 79–82).

Patient characteristics, another major component in the conceptual model, may also influence physicians' reactions to uncertainty and how they convey uncertainty to patients. This influence is described by one respondent. "Some physicians . . . are perturbed by uncertainty. Most patients as well eschew uncertainty. They prefer to polarize, to see the world as black and white without shades of grey. I try to use a minimum of uncertainty in my talks with patients. Most patients prefer that. But I also indicate that uncertainty exists."

Physicians' perception of patients' desires for certainty may affect how uncertainty is handled by physicians and conveyed to patients. Additionally, other patient characteristics may affect physicians. "Other appropriate questions might be: Do you worry about confusing the patient when you decide to discuss uncertainties of his/her diagnosis? Does your patient's education level or levels of understanding affect your decision to discuss uncertainties of his/her illness? . . . Do patients want to know your uncertainty?"

These verbatim comments illustrate practicing physicians' reactions to uncertainty, point to a number of factors influencing physicians, and highlight the complexity of their reactions to uncertainty. Overall, these spontaneous comments focus on two factors in the model, the type of medical problem and patient characteristics, that may influence physicians' reactions to uncertainty. The other factors noted in the model—test and treatment characteristics, physicians' characteristics, and organizational characteristics—were not mentioned spontaneously by

physicians in this study. Although the influences of medical education have been described by Fox and others, the effect of other institutional forces on physicians and physician-patient relations awaits further research.

DISCUSSION

Our study is one of the first systematic quantitative evaluations of practicing physicians' reactions to uncertainty based on a large sample. From it arises a revised, more complex perspective than was held previously about physicians' responses to uncertainty. Physicians, using a self-report format, acknowledge uncertainty and the implications of uncertainty to a greater degree than described by previous qualitative research, which focused on denial of uncertainty and techniques for controlling or minimizing uncertainty (Light 1979; Bosk 1980; Mizrahi 1984) used by trainees. These differences may be due to differences (1) between trainees and practicing physicians, (2) in the predominant sociological forces in training and clinical practice settings, and/or (3) in the research methodologies used in these studies. Research including both trainees and practicing physicians *and* using qualitative and quantitative research methods is necessary for a better understanding of the relationship between our findings using a quantitative method and those of previous studies using qualitative methods.

Physicians' reactions to uncertainty vary by certain demographic and professional characteristics. This study begins to map the patterns of uncertainty and the variations in them that actually exist. Although Fox (1957, pp. 239–41), in her seminal work on physicians and uncertainty, chose to focus on the "average" student, she indicated that significant variability exists among medical students in their responses to uncertainty, particularly those resulting from "imperfect mastery of what is currently known." The research done since (primarily with physicians in residency training) yielded a narrower view of these responses (Light 1979; Bosk 1980; Mizrahi 1984). By not studying physicians in practice, this earlier work might have missed some of the complexity and variability of physicians' responses to uncertainty that are evident in practitioners past the training stage of their career; after all, physicians meet a variety of factors that may counter or reinforce the influences of their residency training (Mizrahi 1986, pp. 136–63). Indeed, residency training and academic medical centers may create organizational forces that override other factors during the training period. To what extent these forces create a lasting effect on physicians' reactions to uncertainty is an important question to be answered by longitudinal studies that follow physicians from training into practice.

Fox (1957, pp. 239–41) suggested that variations in medical students' aptitudes, skills, knowledge, and awareness of their own limitations along with variations in the "experiences through which the student becomes acquainted with the uncertainties of medicine" all contribute to the variability in their responses to uncertainty. Wide variations in the levels of agreement with the majority of scale items by physicians in this study also suggest a far more complicated picture than heretofore described. The physician characteristics described by Fox were combined with four other factors to form the conceptual model for this study (fig. 1). The five groups of factors, and their grounding in a sociocultural context, may account for the complexity and variability in physicians' responses to uncertainty. Multidisciplinary research is needed to understand the effects of these factors, internal and external, on physicians' reactions to uncertainty.

However, one consistent theme seems to emerge from all of these studies of medical students, physicians in residency training, and physicians in practice: the fear of personal inadequacy and failure. This theme was first identified by Fox (1957) and further developed by Mizrahi (1984) in her study of residents in internal medicine training. In addition, fear of personal inadequacy and failure was evident among a portion of the physicians in this study. Others have noted this theme among practicing physicians as well (Feinstein 1985; Nightingale 1987, 1988; Williams et al. 1988). In a follow-up of a sample of residents from her original study, Mizrahi (1986, p. 158) found that those in community practice feared failure, and this was "almost always linked to the clinical diagnosis and treatment" of patients. Although this was not a major dimension of physicians' affective reaction to uncertainty, it may indeed be related to the stress from uncertainty that was identified as a major dimension, and it merits further exploration. This indeed may be the burden professionals and experts take on as our society becomes increasingly specialized and technological, relying on science versus religion or magic to explain uncertainties.

Second, our study exemplifies (1) a methodological approach not previously used to study physicians' reactions to uncertainty and (2) a conceptual approach that attempts to integrate internal processes, findings from psychological work, with external influences and findings from sociological and anthropological work. We have demonstrated how, by building upon a body of . qualitative work, an elusive yet important social phenomenon can be analyzed quantitatively. In addition, we have shown, in the process of integrating quantitative and qualitative findings, how new areas for research can be identified. In particular, stress from uncertainty and reluctance to disclose uncertainty are important dimensions for future research as institutional pressures to contain costs force

doctors surrounded by close monitoring to live without exhaustive testing.

Other examples of research areas suggested by our findings and conceptual framework include: comparative studies done on attitudes, intentions, and actual behaviors of physicians in different settings operating under different organizational and financial constraints and incentives; the impact of being sued on medical decision making and physicians' stress from uncertainty; and the identification of structural or cultural barriers, including the educational process, that may prevent constructive disclosure of uncertainties to colleagues or patients.[6] These studies cannot fail to help medical curriculum specialists reorganize protocols for communicating more effectively with patients (Lipkin et al. 1984) and develop new programs for teaching problem solving and decision making (Bursztajn, et al. 1981; Light 1988) under conditions of uncertainty.

Beyond medicine, research has not even begun to uncover the predictors and effects of uncertainty on scientists, commercial airline pilots, business managers and executives, corporate lawyers, junior faculty and other professionals. Do perceptions of uncertainty among members of these occupations highlight dimensions other than stress and disclosure? What structural and demographic variables affect these professionals' reactions to uncertainty? From what we have learned in this study of one profession, the nature of uncertainty almost undoubtedly will vary widely for each of these other occupations according to technological characteristics, social customs, organizational factors, and financial constraints. It is time to move beyond medicine to develop a comparative sociology of uncertainty. This article might be seen as an early contribution to that effort.

Looking at physicians' experience of uncertainty and their fear of failure in the larger context of anthropological and sociological observations suggests a deeper implication for the nature of the patient's role and the doctor-patient relationship. Scientific medicine seems to have assumed an impossible burden. As Parsons (1951, 1970), Douglas (1966), and Fox (1979) imply, the effort of science and rationality to erase ignorance, eliminate uncertainty, and conquer nature is bound to disappoint, anger, and fail. "By disclaiming that the cause and meaning of illness have anything to do with the supernatural or the inherently mysterious," Fox (1979, p. 510) observes, "modern medicine provides no legitimation for the occurrence of such problems of meaning and no institutionalized way of dealing with them." The emphasis on vigorous treatment and a "can-do" attitude makes it all the worse (Payer 1988). Were doctors

---

[6] It can also identify the latent functions of sustaining client uncertainty (Davis 1960; Waitzkin and Waterman 1974; Light 1979).

and patients to reframe their relationship as an attempt to discover a modicum of pattern and effect what improvements they could in a sea of complex biopsychosocial interactions and uncertainty, expectations would be more realistic and role relations more constructive. Were they, like Douglas and Fox, to regard diagnostic systems as attempts to bring order to a large complex of phenomena and scientific medicine as a symbolic and empirical system for identifying some causal relations between pathogens, symptoms, and interventions, modern medicine would come closer to the sensible balance of the Azande between systems for coping with the problems and uncertainties of life and a score of reasons why witch doctors are sometimes wrong and magic does not always work.

APPENDIX

The Appendix contains the original pool of 61 items used to develop the Physicians' Reactions to Uncertainty (PRU) scales. As a group, the study physicians had consistent responses to the 23 items in part 1 and inconsistent responses to the 38 items in part 2. To assess consistency, item responses were collapsed into dichotomous variables: strongly, moderately, slightly agreeing = agreement; strongly, moderately, slightly disagreeing = disagreement. We then defined a "consistent response" among physicians as items where 75%–100% of the physicians had the same response (either agreement or disagreement with the item). The items labeled stress scale (13 items) and disclosure scale (9 items) were retained in the final Stress from Uncertainty and Reluctance to Disclose Uncertainty to Others scales, respectively.

Part 1: Items Where Study Physicians Had Consistent Responses

75%–100% of Physicians Agree

I am frustrated when I do not know a patient's diagnosis (stress scale).

When physicians are uncertain of a diagnosis, they should share this information with their patients (disclosure scale).

I frequently refer patients to other physicians when I am uncertain of a diagnosis.

I worry that I cannot keep up with the medical literature.

Patients who cannot give a clear description of their problem frustrate me.

When I chose medicine as a career, I was not fully aware of the degree of uncertainty it entailed.

If patients accepted the limitations of biomedical science, a physician's job would be much easier.

I often discuss my uncertainty with patients when I am not sure what is causing their problem.

Even when I make patient-care mistakes, I feel confident of my abilities as a physician.

If a patient requests that a test be done, I will usually do the test.

I often feel guilty when I miss a diagnosis.

I am extremely troubled when I think patients of mine may have a bad outcome because of my care.

75%–100% of Physicians Disagree

I am tolerant of the uncertainties present in patient care (stress scale).

Not being sure of what is best for a patient is one of the most stressful parts of being a physician (stress scale).

The hardest thing to say to patients or their families is, "I don't know" (disclosure scale).

I almost never tell other physicians about diagnoses I have missed (disclosure scale).

I never tell other physicians about patient-care mistakes I have made (disclosure scale).

I am afraid other physicians would doubt my ability if they knew about my patient care mistakes (disclosure scale).

If I share my uncertainties with patients, I will increase the likelihood that I will be sued (disclosure scale).

Missing a diagnosis does not bother me.

I find the uncertainty in medicine challenging.

When I am uncertain of a diagnosis, ordering more tests always relieves my anxiety.

I never think about the consequences of my uncertainty in patient care.

Part 2: Items Where Study Physicians Had Inconsistent Responses

Less than 75% of Physicians Agree

I worry about malpractice when I do not know a patient's diagnosis (stress scale).

I frequently wish I had gone into a specialty or subspecialty that would minimize the uncertainties of patient care (stress scale).

I usually feel anxious when I am not sure of a diagnosis (stress scale).

Uncertainty in patient care makes me uneasy (stress scale).

The uncertainty of patient care often troubles me (stress scale).

I find the uncertainty involved in patient care disconcerting (stress scale).

I am quite comfortable with the uncertainty in patient care (stress scale).

The vastness of the information physicians are expected to know overwhelms me (stress scale).

I fear being held accountable for the limits of my knowledge (stress scale).

I always share my uncertainty with my patients (disclosure scale).

I always feel anxious when I consider a potentially fatal disease in a patient's differential diagnosis.

I would be a much better physician if I read more of the medical literature.

I always feel relieved when a consultant finds something I missed.

I frequently think about the legal implications of missing a diagnosis.

I virtually always feel my clinical ability is adequate.

I have confidence in my knowledge of the current medical literature.

In some ways being unsure of how a patient's condition will turn out is more stressful than knowing for sure that it will turn out badly.

I owe it to my patients to appear confident and knowledgeable, even if I am unsure of what they have or how to treat them.

I frequently order tests to cover all the bases if I am uncertain of a diagnosis.

I always feel frustrated when I do not know a patient's diagnosis.

Confronting one's own limitations in medical practice is a very stressful experience.

Less than 75% of Physicians Disagree

When I am uncertain of a diagnosis, I imagine all sorts of bad scenarios—patient dies, patient sues, etc. (stress scale).

If I do not make a diagnosis, I worry that the referring physician will stop sending patients to me (disclosure scale).

If I shared all of my uncertainties with my patients, they would lose confidence in me (disclosure scale).

If I do not know a patient's diagnosis, I always order more tests.

To maintain my practice, I must have answers for nearly all of the patient-care problems referred to me.

Uncertainty doesn't bother me.

Patients will seek care elsewhere if I do not have answers for them.

When I do not know the answer to a patient's question, I attribute it to the uncertainty inherent in medicine.

I feel anxious when my patients want a second opinion.

I feel threatened by patients who demand answers.

I feel helpless when a patient demands an answer that I do not have.

Patients often want me to share my uncertainty with them.

To reassure patients, I order more tests when I am uncertain of their diagnoses.

I feel inadequate as a physician when I do not know a diagnosis.

Patients demand too many answers from physicians.

I always feel inadequate when one of my patients has a bad outcome.

Rarely am I bothered by the uncertainty of patient care.

REFERENCES

Applegate, William B. 1986. "Physician Management of Patients with Adverse Outcomes." *Archives of Internal Medicine* 146:2249–52.
Atkinson, Paul 1984. "Training for Certainty " *Social Science and Medicine* 19. 949–56.
Bell, David E. 1982. "Regret in Decision Making under Uncertainty " *Operations Research* 30:961–81
Bickel, Janet. 1988. "Women in Medicine: A Status Report." *New England Journal of Medicine* 319:1579–84
Bidwell, Charles E. 1976. "The Structure of Professional Help." Pp. 889–909 in *Explorations in General Theory in Social Science: Essays in Honor of Talcott Parsons,* edited by Jan J. Loubser, Rainer C. Baum, Andrew Effrat, and Victor Meyer Lidz. New York: Free Press.
Bosk, Charles L 1979. *Forgive and Remember: Managing Medical Failure.* Chicago. University of Chicago Press.
———. 1980. "Occupational Rituals in Patient Management." *New England Journal of Medicine* 303:71–76.
———. 1986. "Professional Responsibility and Medical Error." Pp. 460–77 in *Applications of Social Science to Clinical Medicine and Health Policy,* edited by Linda H. Aiken and David Mechanic New Brunswick, N.J : Rutgers University Press
Bowers, C. A 1987. *The Promise of Theory: Education and the Politics of Cultural Change* New York: Teachers College Press.
Bowman, Majorie A., and Deborah I. Allen 1990. *Stress and Women Physicians,* 2d ed. New York. Springer-Verlag.

Bucher, Rue, and Joan G. Stelling. 1977. *Becoming Professional.* Beverly Hills, Calif . Sage.

Budner, Stanley. 1962. "Intolerance of Ambiguity as a Personality Variable." *Journal of Personality* 30:29–50

Bursztajn, Harold, Richard I. Feinbloom, Robert M Hamm, and Archie Brodsky. 1981. *Medical Choices, Medical Changes: How Patients, Families, and Physicians Can Cope with Uncertainty.* New York: Delacorte.

Cattell, Raymond B. 1966. "The Screen Test for the Number of Factors." *Multivariate Behavioral Research* 1 245–76.

Curley, Shawn P., Stephen A. Eraker, and J. Frank Yates. 1984. "An Investigation of Patients' Reactions to Therapeutic Uncertainty." *Medical Decision Making* 4:501–11.

Curley, Shawn P., Mark J. Young, and J. F. Yates. 1989. "Characterizing Physicians' Perceptions of Ambiguity." *Medical Decision Making* 9.116–24

Davis, Fred. 1960. "Uncertainty in Medical Prognosis: Clinical and Functional " *American Journal of Sociology* 66.41–47.

Detsky, Alan S., Steven C. Sticker, Albert G. Mulley, and George E Thibault. 1981. "Prognosis, Survival, and the Expenditure of Hospital Resources for Patients in an Intensive-Care Unit." *New England Journal of Medicine* 305 667–72.

Douglas, Mary. 1966. *Purity and Danger: An Analysis of Concepts of Pollution and Taboo.* London: Penguin.

Douglas, Mary, and Aaron B. Wildavsky 1982 *Risk and Culture: An Essay on the Selection of Technical and Environmental Dangers.* Berkeley and Los Angeles: University of California Press

Eddy, David M. 1984. "Variations in Physician Practice. The Role of Uncertainty." *Health Affairs* 4.74–89.

Eddy, David M., and John Billings. 1988. "The Quality of Medical Evidence· Implications for Quality of Care." *Health Affairs* 7:19–32.

Eisenberg, John M 1979. "Sociologic Influences on Decision Making by Clinicians." *Annals of Internal Medicine* 90:957–64.

Evans-Pritchard, E. E 1937. *Witchcraft, Oracles and Magic among the Azande.* Oxford. Clarendon.

Feinstein, Alvan R 1985. "The 'Chagrin Factor' and Qualitative Decision Analysis." *Archives of Internal Medicine* 145:1257–59

Fox, Renée C. 1957. "Training for Uncertainty." Pp 207–41 in *The Student-Physician: Introductory Studies in the Sociology of Medical Education,* edited by Robert Merton, George C. Reader, and Patricia Kendall. Cambridge, Mass.: Harvard University Press.

——— 1959. *Experiment Perilous: Physicians and Patients Facing the Unknown.* Glencoe, Ill.: Free Press.

———. 1976a "The Sociology of Modern Medical Research." Pp 102–14 in *Asian Medical Systems: A Comparative Study,* edited by Charles Leslie. Berkeley: University of California Press.

———. 1976b. "Medical Evolution." Pp. 773–87 in *Explorations in General Theory in Social Science: Essays in Honor of Talcott Parsons,* edited by Jan J. Loubser, Rainer C. Baum, Andew Effrat, and Victor Meyer Lidz. New York. Free Press

———. 1979. *Essays in Medical Sociology. Journeys into the Field* New York· Wiley.

——— 1980 "The Evolution of Medical Uncertainty." *Milbank Memorial Fund Quarterly* 58·1–49.

———. 1984. " 'It's the Same, But Different': A Sociological Perspective on the Case of the Utah Artificial Heart." Pp. 68–90 in *After Barney Clark: Reflections on the Utah Artificial Heart Program,* edited by Margery W. Shaw. Austin: University of Texas Press.

———— 1989. *The Sociology of Medicine: A Participant Observer's View.* Englewood Cliffs, N.J.: Prentice-Hall.

Fox, Renée C., and Judith P. Swazey 1975. *The Courage to Fail: A Social View of Organ Transplants and Dialysis.* Chicago: University of Chicago Press, 2d rev ed

Fox, Renée C., Judith P. Swazey, and Elizabeth M. Cameron. 1984. "Social and Ethical Problems in the Treatment of End-Stage Renal Disease Patients." Pp. 46–53 in *Controversies in Nephrology and Hypertension,* edited by Robert G. Narins. New York Churchill Livingstone.

Gerrity, Martha S., Robert F. DeVellis, and JoAnne Earp. 1990. "Physicians' Reactions to Uncertainty in Patient Care A New Measure and New Insights." *Medical Care* 28 724–36.

Gutheil, Thomas G., Harold Bursztajn, and Archie Brodsky 1984. "Malpractice Prevention through the Sharing of Uncertainty: Informed Consent and the Therapeutic Alliance." *New England Journal of Medicine* 311:49–51.

Hardison, Joseph E. 1979. "To Be Complete." *New England Journal of Medicine* 300.193–94.

Hershey, John C., and Jonathan Baron. 1987. "Clinical Reasoning and Cognitive Processes." *Medical Decision Making* 7:203–11.

Hogarth, Robin M. 1987. *Judgement and Choice: The Psychology of Decision.* New York Wiley.

Johnson, Clifford G , Jeffrey C Levenkron, Anthony L. Suchman, and Ralph Manchester. 1988. "Does Physician Uncertainty Affect Patient Satisfaction?" *Journal of General Internal Medicine* 3·144–49

Kassirer, Jerome P. 1989. "Our Stubborn Quest for Diagnostic Certainty: A Cause of Excessive Testing." *New England Journal of Medicine* 320:1489–91.

Katz, Jay. 1984. "Why Doctors Don't Disclose Uncertainty." *The Hastings Center Report* 14:35–44.

Kleinbaum, David G., Lawrence L. Kupper, and Keith E. Muller. 1988. *Applied Regression Analysis and Other Multivariable Methods,* 2d ed Boston PWS-Kent.

Light, Donald W. 1979. "Uncertainty and Control in Professional Training." *Journal of Health and Social Behavior* 20:310–22.

————. 1980. *Becoming Psychiatrists: The Professional Transformation of Self* New York Norton.

————. 1988. "Toward a New Sociology of Medical Education." *Journal of Health and Social Behavior* 29·307–322.

Lipkin, Mack, Jr., Timothy E. Quill, and Rudolph J. Napadano. 1984. "The Medical Interview: A Core Curriculum for Residencies in Internal Medicine." *Annals of Internal Medicine* 100 277–84.

Lusted, Lee B. 1984 "Uncertainty and Indecision." *Medical Decision Making* 4.397–99.

Malinowski, Bronislaw. 1948. *Magic, Science and Religion and Other Essays.* Boston: Beacon.

McNeil, Barbara J., Stephen G. Parker, Harold C. Sox, and Amos Tversky. 1982. "On the Elicitation of Preference for Alternative Therapies." *New England Journal of Medicine* 306:1259–62.

Mizrahi, Terry. 1984. "Managing Medical Mistakes: Ideology, Insularity and Accountability among Internists-in-Training." *Social Science and Medicine* 19. 135–46.

————. 1986 *Getting Rid of Patients. Contradictions in the Socialization of Physicians.* New Brunswick, N.J. Rutgers University Press.

Nightingale, Stephen D. 1987. "Risk Preference and Laboratory Test Selection." *Journal of General Internal Medicine* 2:25–28.

————. 1988. "Risk Preference and Admitting Rates of Emergency Room Physicians." *Medical Care* 26·84.

Nunnally, Jum C 1978. *Psychometric Theory*. New York, N.Y.· McGraw-Hill.

Parsons, Talcott. 1951. *The Social System*. Glencoe, Ill. Free Press.

—————. 1970. "Research with Human Subjects and the 'Professional Complex'." Pp. 116–51 in *Experimentation with Human Subjects*, edited by Paul A. Freund. New York: George Braziller.

Payer, Lynn. 1988. *Medicine and Culture: Varieties of Treatment in the United States, England, West Germany, and France*. New York. Henry Holt.

Peterson, Dane K., and Gordon F. Pitz 1988. "Confidence, Uncertainty, and the Use of Information." *Journal of Experimental Psychology* 14:85–92.

SAS Institute. 1985 *Statistical Analysis System User's Guide· Statistics*, version 5.0. Cary, N.C.: SAS Institute.

Slovic, Paul. 1986. "Informing and Educating the Public About Risk." *Risk Analysis* 6:403–15.

—————. 1987. "Perception of Risk " *Science* 236.280–85.

Sorenson, James R. 1974 "Biomedical Innovation, Uncertainty, and Doctor-Patient Interaction." *Journal of Health and Social Behavior* 15 366–74.

Sox, Harold C. 1986. "Probability Theory in the Use of Diagnostic Tests." *Annals of Internal Medicine* 104.60–66.

Spiegel, David A., Robert C. Smolen, and Cindy K Jonas. 1986. "An Examination of the Relationship among Interpersonal Stress, Morale and Academic Performance in Male and Female Medical Students." *Social Science and Medicine* 23:1157–61

Swazey, Judith P., and Renée C. Fox. 1970. "The Clinical Moratorium. A Case Study of Mitral Valve Surgery." Pp. 315–57 in *Experimentations with Human Subjects*, edited by Paul A. Freund. New York: George Braziller.

Tversky, Amos, and Daniel Kahneman. 1981 "The Framing of Decisions and the Psychology of Choice." *Science* 211.453–58.

Vyner, Henry M. 1988. "The Psychological Dimensions of Health Care for Patients Exposed to Radiation and the Other Invisible Environmental Contaminants." *Social Science and Medicine* 27:1097–1103.

Waitzkin, Howard B. 1983. *The Second Sickness: Contradictions of Capitalist Health Care* New York. Free Press.

—————. 1985. "Information Giving in Medical Care." *Journal of Health and Social Behavior* 26.81–101.

Waitzkin, Howard B., and Barbara Waterman 1974. *The Exploitation of Illness in Capitalist Society*. Indianapolis. Bobbs-Merrill

Wennberg, John E., Benjamin A. Barnes, and Michael Zubkoff. 1982 "Professional Uncertainty and the Problem of Supplier-Induced Demand." *Social Science and Medicine* 16:811–24

Williams, Geoffrey, Anthony Suchman, and Henry Herrera. 1988. "Regret" (letter). *Medical Decision Making* 8:310–11

Wolf, Fredric M , Larry D. Gruppen, and John E. Billi. 1985 "Differential Diagnosis and the Competing-Hypotheses Heuristic: A Practical Approach to Judgement under Uncertainty and Bayesian Probability." *Journal of the American Medical Association* 253·2858–62.

Woodward, Christel A., and O. Adams. 1985. "Physician Resources Data Bank: Nature, Distribution and Activities of Canada's Physicians." *Canadian Medical Association Journal* 132 1175–88.

Young, Elizabeth H. 1987. "Relationship of Residents' Emotional Problems, Coping Behavior, and Gender. *Journal of Medical Education* 62:642–50.

Religion, Disability, Depression, and the Timing of Death[1]

Ellen L. Idler
Rutgers University

Stanislav V. Kasl
Yale University

Despite its importance in Durkheim's work, the subject of religion's influence on health and well-being is rarely addressed in contemporary sociological research. This study of elderly persons in New Haven, Connecticut, examines the prospective relationship between religious involvement and several aspects of health status. Results show significant protective effects of public religious involvement against disability among men and women and of private religious involvement against depression among recently disabled men over a three-year period. Religious group membership also protected Christians and Jews against mortality in the month before their respective religious holidays during a six-year period. The article concludes that religious involvement exerts a strong positive effect on the health of the elderly; that this effect varies by religious group and by sex; that the health behaviors, social contacts, and optimistic attitudes of religious group members may explain part but not all of this association; and that several aspects of religious experience, such as participation in ritual and religion's provision of meaning play a role.

DURKHEIM ON RELIGION IN *SUICIDE* AND *ELEMENTARY FORMS OF THE RELIGIOUS LIFE*

When reference is made to Durkheim's interest in the link between religion and individual or group well-being, the source is nearly always *Suicide,* and rarely the later and more mature *Elementary Forms of the*

Religious Life. In *Suicide,* Durkheim ([1897] 1951) considers the life-preserving functions of religion to be indistinguishable from those of the family, of community groups or workers' organizations, or any other small-scale social structure that has the ability to bridge the gap between the private sphere of the individual and the public sphere of the anonymous institutions of mass society. Researchers typically find support for the *Suicide* model: the concrete and material support that can be offered by members of a religious congregation is increasingly being recognized as a social resource, particularly for the elderly and those in poor health (Taylor and Chatters 1986; Pargament 1982; McGuire 1985). Religious groups offer both the "weak ties" and communication capabilities of large social networks and, at the same time, the intimacy of close ties that can easily penetrate the private sphere, though these benefits may vary from one religious group to another (Pescosolido and Georgianna 1989). Religious-group membership may also provide an incentive for rehabilitation and recovery because it offers to otherwise possibly disengaged elderly a public role to return to. But none of these functions is unique to religious congregations; "religion protects . . . because it is a society" (Durkheim 1951, p. 170).

Durkheim's appreciation of the complexity of religious experience apparently increased during the 18 years between *Suicide* and *The Elementary Forms of the Religious Life* (Durkheim [1915] 1965). In his later writings the idea of religious involvement extends beyond simple contact with a social network to include the division of time and space into the sacred and the profane, the use of symbols as collective representations, and the defining structural feature of congregational worship services, that is, ritual, or the manifestation of collective sentiments. The relevance of what in *The Elementary Forms of the Religious Life* has become a "system of ideas" to problems of illness and physical suffering is inherent in the Durkheimian dichotomy of the sacred (the soul) and the profane (the body). The religious impulse is to separate the body, with its physical frailties and limitations, from the soul, with its spark of the divine. Seeing the body as simply a material envelope, the temporary profane location of the soul, could be an increasingly comforting perception to those in deteriorating health. Indeed, the ability to transcend the body is frequently cited as one of the necessary developmental achievements of the second half of life (Peck 1968). Thus any examination of the relevance

of religion to health must consider its particular role in situations of serious illness and suffering, when individuals are restricted in their activities and one might expect a heightened sense of religious introspection.

Ritual observance, another crucial aspect of religious experience not considered in *Suicide,* has particular importance for the lives of religious-group members. Rituals provide a sociotemporal orientation for individuals within the nested circles of the week, the season, and the year; they let people know where they are in time. In fact, Eviatar Zerubavel (1981) argues that, since antiquity, timekeeping has always been a part of the religious sphere of life. The weekly observance of the Sabbath and the yearly observance of the major holidays preserve the collective memories of the participating group; the periodicity of these celebrations reminds the members of their shared past and emphasizes a continuity with the preceding generations that performed these same rituals in the same way.

Thus, ever since Durkheim, the theoretical framework has been at hand, not only for asking the question, "Does religious involvement have some protective effect on human health?" but also for formulating a possible explanation for such an effect.[2] Why should religious affiliation be related to better health levels? Epidemiological studies of the Mormons and the Seventh-Day Adventists, groups that have very low rates of smoking- and alcohol-related diseases (see Jarvis and Northcott 1987), convincingly demonstrate the regulative or social control functions of religion (Umberson 1987). Even religious denominations that do not proscribe drinking, smoking, or eating meat may have the effect of promoting conventional (low health risk) life-styles among their members by frowning upon excess of any kind (Mechanic 1990).

But there is a presumption in this work that more than just differences in health practices are involved in the observed effects on morbidity and mortality. In addition, the impact of social and emotional support, the integrative functions of religious groups, is widely assumed to be operative but remains less clearly demonstrated. Religious-group membership has usually been included in studies of the impact of social networks on mortality, and it has frequently been found to have independent protective effects net of baseline differences in health status and health practices (for a review of these studies see House, Landis, and Umberson [1988]). It is, however, considered, in the tradition of *Suicide,* as one of a number

[2] Despite the apparently substantial literature on religion and physical and mental health (reviewed in Witter et al. 1985; Larson et al. 1986, Jarvis and Northcott 1987, Levin and Schiller 1987), few published studies are methodologically adequate (some notable exceptions are Berkman and Syme 1979, House, Robbins, and Metzner 1982, Schoenbach et al. 1986; Zuckerman, Kasl, and Ostfeld 1984). This area can still be better characterized as faith than as science.

of possible sources of network contacts, all of which are equally "substitutable" in their functions. Thus the limitations of the empirical evidence for a unique association between religion and health are considerable. The linkage through health habits appears to reduce religious practices and beliefs to healthy life-styles, the social-support functions of religious groups may be functionally indistinguishable from those of other groups, and the crude measurement of the religion variable(s) makes it impossible to tell which dimensions of religiousness, if any, are relevant to health.

And yet the theoretical lines of Durkheim's thought hint at a complex and multidimensional relevance of religious involvement to health. The division of the temporal and physical into the sacred and the profane, the periodic observance of rituals, the physical experience of symbols, the provision of meaning in situations of illness and suffering—all are facets of the religious experience that move its conceptualization well beyond the fundamental functions of social integration and regulation. They suggest a broad agenda for innovation in both measurement and analysis before much more of this complicated association can be understood.

The Durkheimian framework within which we are working also allows us to predict that some religious groups should be more protected than others, that is, that Catholics and Jews, with their tightly integrated societies, should derive greater benefits than Protestants. The benefits of religious affiliation may also vary within religious groups, by participation levels, and according to the intensity of religious feelings of individual members. A highly committed Baptist, for example, may be more protected than a lapsed Roman Catholic. Our individual-level data, containing measures of public involvement with the religious group (attendance at services and the number of other congregation members the respondent knows personally) and the intensity of private religious experience (how deeply religious the respondents consider themselves to be and how much strength and comfort they receive from religion), allows us to ask the integration question directly, without using Protestant, Catholic, or Jewish affiliation as a proxy for how tightly the network is knitted.

If religious-group involvement is related to better health, what intervening mechanisms might be at work? First, the regulative functions of religious groups are readily translated as prescribed or proscribed health behaviors; as we have seen, there is some evidence that the better health habits of some religious group members do have an impact on health. Second, the network contacts provided by religious-group membership may confer more of the same benefits that such contacts are also known to provide: membership in a church or temple may expand the network overall, or it may enlarge the pool from which intimate

contacts may be drawn. Third, religious feelings may spring from an individual's more general positive or optimistic outlook on life. These three mechanisms have in common their independence from religion; to the extent that they mediate the association of religion with health, they explain it as something secular. In short, religious-group membership that is conceived exclusively as integration and regulation leaves out that which is peculiarly religious.

On the other hand, to the extent that these intervening mechanisms do not eliminate the association, we are left with a protective effect that would appear to be more uniquely religious. This would necessitate the move from *Suicide* to *The Elementary Forms of the Religious Life,* with its more complex consideration of the facets of religious experience. In particular, it is the discussion of ritual observances in that work that led us to ask if the impact of religion on health might not be intensified during periods of religious holidays. A heightening of the protective effect in the period before and during these ceremonial occasions would suggest that participation in religious rituals forms an important part of this unique religious contribution. Several articles by Phillips and others (Phillips and Feldman 1973; Phillips 1978; Phillips and King 1988; Phillips and Smith 1990), all of which draw on Durkheimian ideas regarding the significance of ceremonial occasions, have examined the distribution of deaths around individually and socially significant events such as birthdays, elections, and some religious holidays. Phillips's well-known finding was that people, especially famous people for whom the event is more widely recognized, are considerably less likely to die in the six months before their birthday than in the five months after. He attributes this "death-dip" before ceremonial occasions to the powerful need of societies to reaffirm, periodically, the collective sentiments that give them their unity and identity and that remake the bonds uniting the individual with the group. These bonds can apparently be so strong that they cause a postponement of physical death and, simultaneously, a delay in the severing of social relations. This analysis thus adds both a new framework for interpreting the general impact of religion on health and another health outcome—the timing of mortality.

Hypotheses

The constraints of our data inevitably force us to accept less theoretically ambitious goals than we would like: specifically, we are able to examine (1) the relative contributions of inner, private religiousness and externalized, public religious practice to health status, (2) the differing importance of these two for Protestants, Catholics, and Jews, and (3) the timing of death for Christians and Jews around their respective religious holidays.

Because we want to treat health as a broad concept, we assess the impact of religion on three separate but interrelated indicators: functional disability, symptoms of depression, and mortality. The analysis of disability is based on the hypothesis that religious involvement will have a broad protective effect against disability. Among those who are not disabled, religious involvement will be associated with a lesser likelihood of decline of functional ability. Among those who are, greater religious involvement will be associated with improving levels of functional ability over the period of the study, to the extent that religious groups offer the elderly tangible resources for recovery and rehabilitation. In our analysis of depression we hypothesize that religion will have protective effects against distress, particularly in groups that are more physically vulnerable. Both the disability and depression analyses also introduce possible intervening variables that could help explain the association: health practices, social contacts, and psychological attitude. These analyses proceed directly from a preliminary cross-sectional analysis (Idler 1987); the findings of that study were that public religious involvement was associated with better functional ability and fewer symptoms of depression among women in 1982 and that disabled men who were more privately religious were less likely to be depressed.

Following the morbidity analyses, we turn to a mortality analysis that examines the timing of death around religious holidays. To test the salience of these holidays for individual–group bonds, we examine the deaths during the month up to and including the holiday and the month after the major Christian and Jewish observances: Christmas, Easter, Yom Kippur, Rosh Hashanah, and Passover, with the expectation that deaths after the holiday will outnumber those that occur before. The timing and celebration of these holidays serve the crucial function of separating social groups from one another; the identity of religious groups is enhanced by their knowledge of and participation in rituals into which others are not initiated (Zerubavel 1982). Thus we hypothesize that we will find both inter- and intragroup differences: (1) the period before Yom Kippur, for example, will show a reduction in deaths among Jews, but not among Christians, and (2) religious holiday effects on the timing of mortality should be greater for the more observant members of religious groups than for the less observant members.

DATA AND METHODS

The methodological requirements for studies that will extend our knowledge in this area are clear. To examine the effects of religious involvement, and not the effects of the doctrinal beliefs of any particular religious group, probability samples should represent communities that are hetero-

geneous demographically and religiously, not religiously homogeneous communities in which groups such as the Mormons, the Seventh-Day Adventists, or the Amish predominate. To avoid the selection effects or "reverse causation" inherent in the interpretation of cross-sectional associations of, for example, attendance at religious services and disability, study designs should be longitudinal, and follow-up periods must be of sufficient length to detect changes in health that occur after the collection of baseline data. Also necessary for the reduction of selection effects are extensive and reliable measures of physical health status at the baseline. To adequately assess the multidimensionality of religiousness, baseline data should include multiple measures of both public/social and private/subjective forms of religious involvement. This issue is particularly important in elderly populations where greater prevalence of chronic conditions may limit participation in worship services, but not affect devotional practices in the home. To permit a broader consideration of the impact of religion on overall health status, outcomes should include measures of mortality, morbidity, and mental health. To isolate the effects of religion from the often-correlated factors of income, race, education, sex, and age, sociodemographic factors should be controlled. Finally, to disentangle the components of religious involvement, measures of health practices, social networks, and attitudes should be introduced so that the influence of more purely religious feelings and practices can be isolated.

Moreover, even with these strict methodological criteria in place there can be no assurance, based on study design alone, that associations of religious involvement at time 1 with health status at time 2 are evidence for cause and effect. The problem of selection effects operating in cross-sectional studies of the association of public religious participation and health levels is obviously a serious one; but even in a follow-up study, individuals who are too ill or disabled to attend services at time 1 are also likely to be ill or disabled at time 2. The solution to this problem is the final methodological criterion: analyses of time 2 health status measures must control for the same measure at time 1; in other words, the study should focus on health status change. Thus the cohort should not be in a steady state where nothing is happening; the sample also needs to be caught early enough so that the effect of religion has not played itself out, and the baseline adjustment for health status wipes out all the previous effects of religion.

Sample

We test these ideas with data from a stratified probability sample (N = 2,812) of elderly people, aged 65 and over, living independently in the

community of New Haven, Connecticut, in 1982 (Cornoni-Huntley et al. 1986). Respondents were selected from three housing strata: publicly supported housing for the elderly, privately-supported housing for the elderly, and elsewhere in the community. Sampling fractions were higher for men and those living in the two types of age-segregated housing; this method increased the proportion of men and low-income elderly in the sample. The overall response rate for the three strata for the initial 1982 interview was 82%. Response rates for the follow-up interviews in 1983, 1984, and 1985 were 96%, 97%, and 93%, respectively, of all surviving members of the original sample.

Respondents were interviewed in their homes in 1982 and were followed continuously (for the purpose of this study, reporting ends in January 1989). At one-year intervals, in 1983 and 1984, respondents were recontacted by telephone and administered a brief interview that focused primarily on disability and new chronic conditions. In 1985, the surviving sample was administered a full follow-up interview. Continuous mortality surveillance has been maintained since the inception of the study; deaths of sample members are noted from obituaries, hospital records, and proxy information and are verified by obtaining death certificates. No sample members have been lost to mortality follow-up. The lengths of the follow-up periods thus differ for the three health outcome measures; in each case we used the longest follow-up period and as many repeated measures as were available.

Description of Measures

The information on religion obtained for each respondent included their religious affiliation (Protestant, Roman Catholic, Jew, other, none), frequency of attendance at services (1 = never to 6 = more than once a week), the number of other congregation members known (1 = none to 4 = almost all), self-assessments of depth of religiousness (1 = not at all religious to 4 = deeply religious), and of how much strength and comfort they receive from their religion (1 = none to 3 = a great deal). Because we wanted to differentiate the public, social aspects of religious involvement from the private, reflective aspects, we created two two-item indices, one that sums attendance at services and other congregation members known (Cronbach's α = .63) and the other that sums self-reports of religiousness and receiving strength and comfort from religion (Cronbach's α = .71). New Haven is a relatively heterogeneous community; 53% are Roman Catholic (many of them ethnic Italian), 29% are Protestant, of which about half are white and half black, and 13.5%

are Jewish. Unfortunately, no details on Protestant denominations or branches of Judaism are available; a racial breakdown for the Protestants, however, does serve the important purpose of differentiating the generally highly involved blacks from the generally less-involved whites. It is overall a religiously observant and devout community: 41% of this elderly sample attend services at least once a week, and 38% consider themselves deeply religious.

Functional disability, the first of the three outcome measures, is a five-level Guttman scale composed of three subscales tapping a range of "activities of daily living" and physical performance items (for details of construction see Berkman et al. [1986]). These items were asked of respondents at each of the face-to-face and telephone interviews; thus we have identical measures for respondents in 1982, 1983, 1984, and 1985. In 1982, 34% of the sample reported no major or minor disability, while 13% could be considered severely disabled; by 1985 disability levels had increased noticeably: only 28% were without disability and 21% were severely disabled. Of course, the most severely disabled in 1982 were much more likely to die in the 1982–85 period.

The second outcome measure, symptoms of depression, is assessed by the Center for Epidemiologic Studies Depression Scale (CES-D; see Berkman et al. 1986) and is available only for 1982 and 1985. Depression scores changed very little over the three years. The Pearson correlation coefficient for the pair is .47; mean scores remained the same and fewer than 15% scored above the customary diagnostic cutoff point of 16.

The third outcome measure, mortality from all causes, is based on ongoing surveillance of the sample and mortality data are available from 1982 through early 1989. In an initial sample of 2,812, 1,037 deaths had occurred by the end of January 1989, nearly 37%. All but two of these deaths are identified by their exact date.

The most important confounding variable for this analysis, besides the baseline measures of disability and depression, is the general physical health status of the respondent at baseline. We had available to us in the baseline data self-reported diagnoses of 10 chronic conditions (myocardial infarction, stroke, cancer, diabetes, cirrhosis, broken hip, other broken bones, arthritis, Parkinson's disease, amputation); three standard symptom scales for the diagnosis of (1) angina (Rose 1962), (2) intermittent claudication (Rose 1962), and (3) bronchitis (Higgins and Keller 1973); a history of hospital or nursing home admission in the past year; systolic and diastolic blood pressure as measured by the trained interviewer; and the total number of different prescription medications the respondent had taken in the previous two weeks. All of these measures were included in the analysis as controls. In addition, the previously described measure

for functional disability and a measure of change in disability from 1982 to 1985 were included in the depression models.[3]

We also include several groups of other potential confounding and intervening variables. Control variables include demographic characteristics of age and education, sex, race, and perceived income adequacy.[4] Our intervening variables contain a set of health practices, including smoking, alcohol consumption, exercise, and a body-mass index, three indicators of social networks (the number of people seen frequently, the number of close friends and family members, and a dummy variable for being married), and an indication for an attitude of optimism (Cronbach's $\alpha = .56$) (for overall sample characteristics see Cornoni-Huntley et al. 1986).[5]

Statistical Methods

All analyses were performed with the total sample and also separately by sex. In most cases the results differed little for men and women, and we present total sample analyses that control for sex. In the instances where they did differ, we present sex-specific results. Our analyses were performed sequentially, beginning with models containing just the religion and physical health status measures, then adding sociodemographic variables, then three sets of possible mediating variables (health practices, social networks, and optimism), and finally the baseline measure of the outcome variable. In addition, we tested interactions of the two

[3] Our variables, then, were (1) chronic conditions (respondent scores 1 for each of 10 conditions reported), (2) hospitalized in 1981–82 (1 = yes, 0 = no); (3) nursing home in 1981–82 (1 = yes, 0 = no); (4) current medications (score "1" for each prescription observed by interviewer), (5) pain symptoms (add "1" each for angina, intermittent claudication, and chronic cough as diagnosed by symptom scales), and (6) high blood pressure (1 = systolic pressure over 140 or diastolic pressure over 90, 0 = otherwise)

[4] The coding for these variables was. sex (1 = male, 2 = female), age (in years); education (in years), income (1 = not enough money to make ends meet, 0 = otherwise); race (1 = white, 0 = otherwise).

[5] These variables were coded present smoker (1 = yes, 0 = otherwise); past smoker (1 = yes, 0 = otherwise), alcohol consumption (1 = mean + 1 SD for frequency of drinking per month × amount usually consumed × alcohol content; 0 = otherwise); exercise (sum of 0 = never, through 2 = often for active sports or swimming, taking walks, and doing physical exercise); body-mass index (1 = mean + 1 SD for weight/ [height²] [kg/m²]; 0 = otherwise); social contacts (sum of number of children, friends, relatives, and neighbors that respondent sees frequently); intimacy (sum of number of children, friends, and relatives respondent feels very close to), marital status (1 = married, 0 = otherwise); and optimism (sum of 1 = rarely, through 4 = most or all of the time for two questions. "Even though I don't always understand why things happen, I have faith that they will turn out all right" and "I have always had a sense that I belong, that I was a part of things").

religiousness variables with the score for 1982–85 change in functional ability in the analysis of depressive symptoms.

Because of the complex, stratified, and clustered nature of the sample design, the assumption of simple random sampling necessary for ordinary least squares (OLS) regression was inappropriate. It was necessary to incorporate a weighting factor that would correct for the different sampling fractions used for the three strata's males and females, to make the sample approximate the 1980 census data for New Haven, and also to correct for differential nonresponse. The resulting (unstandardized) regression coefficients then had to be tested for significance with standard errors that had been adjusted (usually upward) by an application of the Taylor series for estimating variances of linear functions (Lee, Forthofer, and Lorimor 1989). The Taylor series procedure used for our analyses of disability and depression is SURREGR (Holt 1977), which in most cases increases standard errors and thus results in stricter tests of significance.

For our analysis of the timing of mortality, we performed binomial tests on the frequency of deaths in the 30 days before and including the holiday and the 30 days after it. Based on a null hypothesis of no differences, the probability would be equal ($P = 0.5$) that a death would occur in either period. The normal curve approximation is applicable when the expected frequency is greater than five ($nP \geq 5$), hence the z test (with continuity correction) was appropriate whenever the number of deaths in the two months was 10 or more, as was true in nearly every case. No controls for covariables were necessary as the sample serves as its own control in the before and after periods, but analyses were performed on various subgroups, including males and females, Protestants and Catholics, and the more and less observant (in terms of attendance at services). Four of the five holidays (Easter, Yom Kippur, Rosh Hashanah, and Passover, but not Christmas) move around the calendar; the exact dates of the five holidays in each of the six years 1982–88 were obtained from the *Information Please Almanac, Atlas and Yearbook* (1981).

RESULTS

Religion and Functional Disability: Resources for Recovery

We have three follow-up measurements of functional disability available to us, in addition to the baseline measure. Table 1 reports the regressions of 1983, 1984, and 1985 functional disability on religious affiliation, public and private religiousness, 1982 functional disability and physical health status, sociodemographic controls, and the three sets of mediating variables. We present only the full models because little additional information was gained from the sequential, progressively elaborated models.

For public religiousness, the results are very similar across the three years. There is a highly significant inverse association of 1982 public participation in religion with functional disability in 1983, 1984, and 1985, even when 1982 levels of disability are held constant. It appears that people who were highly involved with their church or synagogue in 1982, regardless of their level of functional ability, were likely to be less disabled one, two, and three years later. Note that there are no main effects of being Protestant, Catholic, or Jewish.

At the same time, we see a significant positive association of private religiousness with 1984 (but not 1983 or 1985) disability: high levels of private religiousness (being deeply religious and getting strength and comfort from religion) in 1982 appear paradoxically to be associated with poorer outcomes in functional ability at two, but not three years. The major difference among these three models, besides the length of follow-up, is the number of respondents remaining in the sample. This attrition, of more than 100 between 1983 and 1984, and more than 150 between 1984 and 1985, is almost exclusively due to mortality. If the attrition due to mortality came primarily from the less privately religious respondents, then what we are seeing would actually be a protective effect of religion because the more religious respondents would be the survivors. On the other hand, if the deaths came from the more religious then it would appear that those individuals who reported receiving a great deal of strength and comfort from religion were in fact approaching death, perhaps with some awareness.

In order to test these explanations, we ran the 1984 model using just the sample of 1985 survivors: the association remained significant, and thus was not explained by the attrition of the more privately religious. This persisting association of higher levels of private religiousness in 1982 with greater disability in 1984 is evidently a survivor effect; in other words, the deaths in the 1982–84 period must have come from the less privately religious or the association would have disappeared. Thus what looks at first glance like a tendency for the more privately religious to become more disabled is actually a short-term protective effect of private religiousness against mortality, but not against disability.

But then the association completely disappears by 1985, and we are left with a longer-term absence of effect. Following the reasoning above, the additional 1984–85 deaths must apparently have come from the highly religious group, those who disproportionately survived until 1984. Thus the diminished sample again resembles the baseline sample with respect to this association. The fluctuation in the effects makes them difficult to summarize neatly, but they do make the point that the mortality in the sample is an important consideration in understanding the relationship of private religiousness to morbidity.

TABLE 1

RELIGIOUS INVOLVEMENT AND FUNCTIONAL DISABILITY IN THE YALE HEALTH AND AGING PROJECT, 1982–85*

	1983 Disability (N = 1,822)		1984 Disability (N = 1,714)		1985 Disability (N = 1,546)	
	b	P	b	P	b	P
Religion.						
Public religiousness	−.060	.000	− 048	004	− 060	003
Private religiousness	024	.549	092	005	004	899
Protestant	− 064	565	074	.421	.090	447
Jewish	−.109	.338	103	.365	− 018	.830
Other	− 293	238	− 022	906	−.194	.256
Health status						
Chronic conditions	.101	015	100	005	136	001
Hospitalized in last year	214	003	059	580	− 098	394
Nursing home in last year	.225	235	216	232	− 228	.303
Current medications	058	.002	.087	000	076	000
Pain	.091	.345	.139	.203	007	.947
High blood pressure	− 158	013	−.154	.071	−.170	065

Sociodemographics.						
Female	.041	.543	157	009	.224	001
Age	032	000	046	000	041	000
Education	-.001	.920	-020	042	-.027	003
Low income	.082	574	113	406	-027	876
White	-149	.269	-044	751	.038	.781
Health behaviors						
Overweight	.335	.002	213	.049	195	.024
Exercise	-167	000	-092	.054	-181	000
High alcohol use	-.018	.877	-119	417	-096	.444
Present smoker	-016	870	072	500	-124	292
Past smoker	.026	.774	162	054	-021	.852
Social networks.						
No. of social contacts	014	023	008	281	010	.076
Close friends and relatives	-018	003	-012	046	000	.970
Married	-052	492	-023	776	-021	.791
Attitude.						
Optimism	-.039	006	-065	001	-.039	.026
Baseline control:						
1982 functional disability	421	.000	417	UUU	.41U	UUU
R^2		39		43		.40

* Significance tests for unstandardized regression coefficients are based on standard errors adjusted by Taylor series method for complex sample design

Other factors besides religiousness, physical health, and 1982 functional disability that contribute significantly to the prediction of 1985 disability are being female, being older, and having less education. Several variables from the sets of possible intervening variables were also significant: exercise, being overweight, number of social contacts (the more disabled had more), and optimism. From the sequential analyses, however, it was clear that none of these variables altered the association of public religiousness with disability, thus we can say that neither health behaviors, nor social networks, nor general attitudes help explain why religiousness should be associated with better functional ability. Public involvement with a religious community thus appears to have some substantial, independent benefits for respondents' ability to perform activities of daily living.

But are the effects the same for Catholics, Protestants, and Jews? Although we saw no main effects of these religious groups on disability, the impact of religiousness might be stronger for one group than for another. To test these interactions, we ran the entire sequence of 1985 models separately for Catholics, Jews, and white and black Protestants, and we found some substantial differences. The effect of public religiousness became stronger for Catholics ($b = -.071$, $P = .0015$) than it had been for the entire sample. An initially strong association for Jews was reduced to marginal significance ($b = -.085$, $P = .068$) by the introduction of the health behavior variables. A marginal association for black Protestants disappeared with the introduction of the health behaviors, and none was ever present for the white Protestants. The protective effects of attendance at religious services are thus unevenly distributed, with the benefits going primarily to Catholics, and to some extent to Jews, very few to black Protestants, and none at all to white Protestants. Also, the healthier lifestyles of the more observant Jews and black Protestants may play a role in explaining their better functional ability.

The findings regarding public religious involvement also raise a question of the nature of the effect. When we control for disability level in 1982, and then predict 1985 disability, the significant association of public religious involvement with the outcome could be primarily due to (1) more recovery among the initially disabled (a rehabilitation hypothesis), (2) less likelihood of decline in functional disability (a prevention hypothesis), or (3) both. To test these alternatives, we ran design-effect-adjusted logistic regression models (Shah 1984) with outcome variables that dichotomized functional disability into (1) improved ($N = 287$) versus same or worse ($N = 1,211$) and (2) worse ($N = 978$) versus same or improved ($N = 520$). The results (not shown here) demonstrated that public religious involvement is somewhat more strongly associated with improvements

in functional ability ($b = .149$, $P = .0014$), than it is with prevention of decline ($b = -.080$, $P = .0647$), although both are clearly involved.

Public involvement with a religious community, especially a Catholic or Jewish community, thus appears to have some substantial benefits for respondents' ability to perform activities of daily living, both because it promotes rehabilitation and return to functioning and because it prevents further deterioration.

Religion and Depression: Moderating Vulnerability

The results of the analysis of depression symptoms are shown in table 2. Because the hierarchical approach yields no additional information, we present only the full model for the total sample and the results of a model containing an interaction term for males only. Neither public nor private religiousness has an impact on 1985 depression when 1982 depression is accounted for. There is, however, a main effect of religious preference: Jews were considerably less likely than Catholics to become depressed between 1982 and 1985. Few other variables are predictive: the most important are functional disability in 1982 and change in disability from 1982–85, but high alcohol use and being married in 1982 are also significant or marginally so. Married status in 1982 is positively associated with depression because this group is at risk of bereavement. The most notable aspect of this full model is that very little of the change in depression symptoms over the period is explained; the 1982 and 1985 measures were correlated at .61. The entire sequence of models was run for the four religious groups separately. There were no effects of public or private religiousness for Catholics, Jews, or black or white Protestants.

Because we were particularly interested in the impact of religiousness on those who experienced new health problems, we continued the analysis of depression by examining interactions between religiousness and change in disability level. These analyses were sex specific; the significant results are reported in the final columns of table 2. Because of the much smaller size of the male sample, we found it necessary to reduce missing data by including in these models only the health status controls and those variables that were significant predictors of 1985 depression for men. The model shows that private religiousness protects men who have worsening disability from becoming depressed in 1985, even with adjustments for their level of depression in 1982. These findings gain interpretive appeal from the disability findings presented above: private religiousness, especially in the sense of getting a great deal of strength and comfort from religion, can easily be seen as responsive to trouble, in this case the recent trouble of an increase in functional disability. We conclude that religious involvement provides some protection against depres-

TABLE 2

RELIGIOUS INVOLVEMENT AND SYMPTOMS OF DEPRESSION IN THE YALE HEALTH AND
AGING PROJECT, 1982–85*

	FULL MODEL, FULL SAMPLE (N = 1,447)		INTERACTION MODEL, MALES ONLY (N = 603)	
	b	P	b	P
Religion:				
Public religiousness	−.181	.195	− 168	.095
Private religiousness	.105	.682	115	.653
Protestant	.327	638		
Jewish	− 1.679	.020		
Other	1.316	.449		
Health status:				
Chronic conditions	−.086	.724	.001	.999
Hospitalized in last year	− 306	598	−.811	273
Nursing home in last year	.028	972	4.486	.056
Current medications	.020	.898	−.038	.805
Pain	404	.564	−.938	284
High blood pressure	− 229	625	− 189	.800
Functional disability	1.239	.000	1 107	.002
Change in disability, 1982–85	1.233	000	4.088	.004
Private religion × change in disability			−.501	.029
Sociodemographics:				
Female	.063	.892		
Age	.002	.957	−.101	.057
Education	−.084	242		
Low income	−.264	.797		
White	478	.624		
Health behaviors.				
Overweight	−.317	.637		
Exercise	153	.587		
High alcohol use	1.227	.069		
Present smoker	− 019	.973		
Past smoker	− 529	.281		
Social networks·				
No of social contacts	− 025	.474	−.056	.024
Close friends and relatives	− 023	.547		
Married	1 859	.001		
Attitude·				
Optimism	−.236	.148		
Baseline control:				
1982 CES-D score	.474	.000	.518	.000
R^2		.36		.37

* Significance tests for unstandardized regression coefficients are based on standard errors adjusted by Taylor series method for complex sample design.

sion for the elderly, especially the vulnerable group of elderly men with deteriorating functional status.

Religion and Mortality: The Postponement of Death from Sacred to Profane Time

Before discussing the analysis of the timing of deaths around religious holidays, we should briefly report the results of the overall mortality analysis that we performed. We estimated logistic regression models (Shah 1984) for all-cause mortality through January 1989 for men and women separately. There was no association of public or private religiousness with mortality for men when health status in 1982 was controlled. With only health status controlled, there was a reduced risk of mortality for women, but this association disappeared when sociodemographic factors (age in particular) were controlled. In addition, we tested interactions between the two forms of religiousness and physical health status (Zuckerman et al. 1984) but again found no associations.

We then performed another type of mortality analysis that appeared to have great relevance for religion, that is, an examination of the timing of death around religious holidays. Tables 3 and 4 show the results of our analysis of deaths in the 30 days up to and including the day itself (-30) and the 30 days after the holiday ($+30$).[6] Table 3 presents frequencies for the number of deaths around Christian holidays, and the standard errors and P-values from the binomial z tests for the two periods, among both Christians and Jews. The results are very strong. Among Christians, with the exception of the less observant white Protestants at Easter, there are fewer deaths during the holiday and the 29 days preceding it than during the 30 days after. This is true not only for the total sample and for both holidays combined, but also for Christmas and Easter separately, for males and females, and for less and more observant Christians, with the exception noted above. (Because we were interested specifically in ritual participation, as opposed to religious feelings, we grouped more and less observant people on the basis of their usual frequency of attendance at religious services; those who attended once a month or more often were considered "more observant," the others "less observant.") These differences between pre- and postholiday deaths produce statistically significant z scores (one-tailed tests with continuity correction) for males and females, for the more and less observant, for Catholics and the less observant black Protestants (but not for white Protestants), at

[6] We include deaths from all causes Only 11 of the 1,037 deaths in the sample were from external causes; removing the eight that occurred around a holiday from the analysis had no effect on the results.

TABLE 3

Deaths among Christian and Jewish Yale Health and Aging Project Respondents
Thirty Days before and after Christian Holidays, 1982–88*

	Christmas				Easter				All Christian Holidays			
	−30	+30	SE	P	−30	+30	SE	P	−30	+30	SE	P
Among Christians:												
Males	25	38	3.97	.039	33	42	4.33	125	58	80	5.87	.025
Females	30	51	4.50	.007	37	40	4.39	326	67	91	6.28	.024
More observant	25	41	4.06	.018	32	40	4.24	145	57	81	5.87	.017
Less observant	30	48	4.42	.016	38	42	4.48	288	68	90	6.28	.034
Catholics	35	62	4.92	.002	42	56	4.95	066	77	118	6.98	.001
More observant	17	30	3.42	.021	23	32	3.71	089	40	62	5.04	.011
Less observant	18	32	3.53	.017	19	24	3.28	181	37	56	4.82	.019
White Protestants	9	11	2.23	.251	17	10	2.59		26	22	3.46	
More observant	2	3	1.11	.187	2	3	1.11	187	4	6	1.58	.171
Less observant	7	8	1.93	.302	15	7	2.34		22	15	3.04	
Black Protestants	10	16	2.54	.085	11	14	2.50	212	21	30	3.57	.081
More observant	6	8	1.87	.212	6	5	1.65		12	13	2.50	.345
Less observant	4	8	1.73	.075	5	9	1.87	092	9	17	2.55	.039
For all Christians	55	89	6.00	.002	70	82	6.16	147	125	171	8.59	004
Among Jews:												
Males	6	6	1.73	.386	5	9	1.87	090	11	15	2.55	.164
Females	7	10	2.06	.166	10	5	1.94		17	15	2.82	
More observant	5	5	1.58	.374	4	7	1.65	113	9	12	2.29	.192
Less observant	8	11	2.18	.179	11	7	2.12		19	18	3.04	
For all Jews	13	16	2.69	.230	15	14	2.69		28	30	3.81	.348

* P values are calculated from one-tailed z tests with continuity correction

Christmas and Easter combined and Christmas alone, but not at Easter. (Because we were interested only in the one-tailed test, standard errors but no P-values are reported for tests in which there were more deaths before the holiday than after it.)

For Jews, on the other hand, there are almost as many cases where there are more deaths before the Christian holiday than there are after it. In no case are there significantly fewer deaths of Jews before the Christian holiday than after. Thus we confirm the hypothesis of intergroup differences. However, we find no differences between the less and the more observant Christians in their timing of death around Christmas, except in the case of the black Protestants among whom only the less observant show a significant pattern, so our other hypothesis, that of intragroup differences is not supported. It may be that the celebration of Christmas is so widespread in American culture that even (perhaps *especially*) the normally nonobservant Christians are drawn to anticipating it.

Table 4 shows the results of the analysis for the Jewish holidays of Passover, Rosh Hashanah, and Yom Kippur. The first notable difference between these results and the previous ones is that the patterns differ for Jewish males and females. Jewish males show the expected pattern of fewer deaths before every holiday, and the differences for before- and after-holiday deaths are significant for Passover, Yom Kippur, and all Jewish holidays. However, Jewish females show no such differences, and, in fact, we see more deaths among Jewish females *before* than after the holidays. While our significance testing reflected a hypothesis that did not anticipate this finding, we note that the finding is very similar to that made by Phillips and King (1988). When they analyzed their sample of 1966–84 California deaths separately by sex, they found a "Passover pattern" in deaths for Jewish males that was not present at all among the Jewish females. Their interpretation of this finding is that Passover is less religiously and socially significant for Jewish females because women are excluded from any role in the important aspects of the ritual, such as the recitation of the Exodus story, the asking of the four questions, and the washing of the hands. Thus, they argue, Jewish men have a greater psychosocial investment in the observance of this holiday (see also Zerubavel 1981, p. 134). Our data agree with theirs and, in fact, extend the finding to at least Yom Kippur, if not Rosh Hashanah.[7]

---

[7] The close timing of Rosh Hashanah and Yom Kippur (Rosh Hashanah always occurs nine days before Yom Kippur) led us to make a Bonferroni adjustment for multiple comparisons. Setting $\alpha = .025$ would make the difference for males at Yom Kippur nonsignificant More important, however, is the fact that there is no before-after difference around Rosh Hashanah, the first and least important of the two holidays.

TABLE 4

DEATHS AMONG JEWISH AND CHRISTIAN YALE HEALTH AND AGING PROJECT RESPONDENTS THIRTY DAYS BEFORE AND AFTER JEWISH HOLIDAYS, 1982–88*

	Passover				Rosh Hashanah				Yom Kippur				All Jewish Holidays			
	-30	+30	SE	P	-30	+30	SE	P	-30	+30	SE	P	-30	+30	SE	P
Among Jews:																
Males ..	5	11	2 00	040	6	7	1.80	291	3	8	1 66	035	14	26	3 16	.020
Females ..	9	4	1 80		8	3	1 66		7	3	1.58		24	10	2.91	
Less observant ...	11	7	2 12	035	11	7	2 12	341	9	6	1 94	021	31	20	3 57	.019
More observant .	3	8	1.66		3	3	1.22		1	5	1 22		7	16	2.40	
For all Jews ..	14	15	2 69	.359	14	10	2 45		10	11	2.29	330	38	36	4.30	
Among Christians																
Males	35	32	4.09		32	27	3 84		34	24	3 81		101	83	6 78	
Females	38	34	4 24		32	31	3 97		31	37	4.12	198	101	102	7.12	.444
Less observant ..	42	36	4.42		34	27	3.90	.397	33	31	4 00		109	94	7 12	
More observant ...	31	30	3 90		30	31	3.90		32	30	3 94		93	91	6.78	
Catholics	43	46	4.72	337	43	43	4 63	456	47	39	4 63	166	133	128	8 08	.401
White Protestants ...	16	9	2 50		9	6	1 94		7	10	2 06		32	25	3.77	
Black Protestants	15	11	2.54		12	9	2 29		11	9	2.24		38	29	4 09	
For all Christians	73	66	5 87		64	58	5 52		65	61	5 61		202	185	9 84	

* P values are calculated from one-tailed z tests with continuity correction

Our hypothesis of intragroup differences by regular observance level is also confirmed. The reduction in deaths before holidays occurs only for the more observant and not at all for the less observant Jews. In fact, as did the females, the less observant had higher frequencies of deaths before the holidays than after. This coincidence of findings led us to examine the "more observant" category more closely; in fact, this category is made up mostly of males (16 of 23). While on the one hand this means a confounding of maleness with the tendency to be observant, thus making our separate analyses somewhat overlapping, it also provides support for the Phillips and King (1988) interpretation, that the higher levels of ritual involvement among Jewish men do indeed provide them greater benefits.

The intergroup differences in the table are less complicated. In nearly every case, the number of deaths among Christians in the preholiday period matches or exceeds the number in the postholiday period. Thus we find that the reduction in deaths before Jewish holidays occurs only among Jews, particularly among the more observant males.[8]

---

the anticipation of Yom Kippur apparently prevents an increase in deaths immediately after Rosh Hashanah Along these same lines, our confirmation of the hypothesis of intergroup differences despite the close timing of Passover and Easter strengthens rather than weakens our argument.

[8] Several alternative interpretations of these findings are available in Phillips and Feldman (1973). It is possible that the registration process for deaths is disrupted by major holidays; we consider this to be an unlikely explanation for our findings because so many of the deaths in our sample took place in the hospital They also consider the effect of seasonal variation in mortality; a long-term upward trend in mortality could favorably bias the findings. Our analysis of not one, but five religious holidays at different times of the year makes this explanation unlikely, but we examined the six-month period centered on each of the five holidays in each of the six years for such a trend. We found no trend at all for the autumn holidays of Rosh Hashanah and Yom Kippur; we found an upward trend in mortality for the winter holiday of Christmas and a downward trend for the spring holidays of Easter and Passover. The upward trend is not the sole explanation for the Christmas results, for one thing, we see the pattern only among Christians, moreover, the dip before the holiday is lower than the previous overall lower level in mortality, and the peak afterward is higher than the subsequent overall high. For the spring holidays, the downward trend only underscores the importance of our findings for Passover and Easter The upward and downward trends for the two Christian holidays counterbalance each other in the significant findings for "all Christian holidays", the only trend around a Jewish holiday is a downward one. Thus our findings are significant despite the trends, not because of them Another explanation Phillips and Feldman (1973) propose is that unusual excitement before the holiday may actually cause the death after it; that is, the before-after difference may be due primarily to extra deaths after the holiday, rather than fewer before. To examine this issue, we compared the average number of pre- and postholiday deaths to the other months of the year; this comparison showed an absence of "death peaks" after the holidays. The differences are due to a reduction in deaths before the holiday, not an increase after it Finally, one should consider the effects of medical care perhaps physicians make extraordinary efforts to keep people

DISCUSSION

These findings suggest that religion does play a profound and rather complicated role in the lives and health of the elderly people in this sample and that many of the elements in Durkheim's formulations appear quite relevant. Our large and representative sample, multiple measures of both religion and health status, prospective study design, conservative techniques for the analysis of data from stratified samples, and the range of control variables considered suggest that the observed effects of religion in this study are not due to selection or to inadequate controls for other confounding factors. Rather than dwelling on these methodological issues, however, we wish, in this concluding section, to address the theoretical import of the pattern that our findings take.

Public religious involvement, meaning frequent attendance at religious services and active participation in the social life of the congregation, was strongly associated with improvements in functional ability levels and the capacity to perform basic self-care and more difficult physical performance activities. It was also associated with change in depressive symptoms for functionally disabled men over the three-year period. Our study design permitted a replication of the functional disability analysis for variable follow-up intervals, and the findings indicate that the effects of public religious involvement observed after one year last for a period of at least three years. From an epidemiology of aging standpoint, this is a meaningful finding for several reasons. Functional disability is a central indicator of health status in the elderly; in this sample it is strongly associated with both mortality (Idler, Kasl, and Lemke 1990) and depression (Berkman et al. 1986). Functional disability is also the basis for projections of active life expectancy, or remaining years of independent living (Katz et al. 1983).

Our measure of public religiousness was intended to tap the congregational dimension of religious involvement, the "obscure but intimate relations" (Durkheim 1965, p. 257) of this society, and the religious experience that takes place in its midst. Our sequential models (not shown) indicated that the potential intervening variables examined, including health practices, numbers and closeness of social contacts, and optimism, did not account for any portion of the association between public religiousness and reduced disability. Thus the findings suggest that it is not behaviors, networks, or attitudes alone, but some other element of the religious experience felt by those who attend services frequently that

---

alive in the period before important events. This explanation was dismissed by Phillips and Feldman because, in their historical data, there was as much evidence of death dips before birthdays in the period before 1894, when medical care was ineffective at best, as there was after.

makes the difference. Future research should perhaps focus on the spiritual supports churches and synagogues offer for healing, recovery, and rehabilitation or on more multidimensional measures of the public, congregational, and ritual aspects of religious involvement. The underlying mechanisms of this support await clarification; nevertheless, because it is in public and not private religious involvement that we see these effects, we feel we can concur with Durkheim (1951, p. 170) that "the beneficent influence of religion" in the case of functional disability is due to its capability of "supporting a sufficiently intense collective life," or, in other words, that "it is a society."

The findings regarding the self-assessments of being deeply religious and receiving a great deal of strength and comfort from religion are more complicated: private religiousness shows an opposite association with disability than does public religiousness, and depression does not show the same association with private religiousness that is seen for disability. Furthermore, the association with depression is sex specific. The only easy conclusion is that it is, indeed, important to have separate measures for different components of religious involvement. Beyond that, we can only offer a few speculations.

Regarding the association of greater private religiousness with poorer near-future prognosis for level of functioning, at least two areas of speculation suggest themselves. One is that there is an acute reactive component of private religiousness that is influenced by poor health status; in effect, people who say they receive "a great deal" of strength and comfort from religion are probably people who have a reason for needing it. The fact that this finding does not persist among the 1985 survivors indicates that the association was strongest among the portion of the sample whose disability was an (accurate) indication of the poorest prognoses. The difficulty with this speculation is that the analyses adjusted for baseline poor health and disability. Thus, the speculation is viable only if the reactive component of private religiousness is sensitive to deteriorating health that is not yet fully reflected in the health status and disability measures at baseline. A second area of speculation is that private religiousness indicates a turning inward so that the goal is the spirit's inner peace rather than the body's social role performance and social interaction with others. The turning inward could undermine motivation for recovery or for the maintenance of good functioning.

With respect to the role of private religiousness as a buffer in the effect of new disability on depression, the finding is almost a face validation (but a prospective one) of the question on private religiousness as a source of comfort. What is hard to explain is that the effect is seen for men only. This sex-specific finding reiterates cross-sectional findings in these data (Idler 1987) and others (Siegel and Kuykendall 1990). Possible explana-

tions for this finding include sex differences in the conceptualization and measurement of depression, a differential importance of troubles (disability or bereavement) for men and women, or differential interpretations of the meaning of private religiousness.

Our third finding concerns the ritual element in the Durkheimian conception of religion. We found that the timing of mortality among elderly Christians and Jews was closely linked to their own religious holidays, when the most elaborate ritual preparations and enactments are made. One might interpret this finding in two ways: these times are special either because of their heightened spiritual or their heightened social significance. Their spiritual significance derives from the performance of rituals and the presence and manipulation of symbols that are seen infrequently but are central to key historical events in the faith. The celebration of Communion on the Thursday preceding Easter commemorates the Last Supper, at which Christ transformed the meaning of bread and wine into body and blood. Likewise, the unleavened bread and bitter herbs of the Passover seder commemorate the slaughter of the firstborn and the flight of the Jews from Egypt. The Durkheimian argument is that religious symbols hold their special power through a concentration of meaning. They are even more powerful than the things they are material expressions of, because they can more easily be held in the mind, especially at times when the individual is separated from the group (1965). But religious symbols such as these also involve the body and may derive additional power from their evocation of physical sensations, which provides an underpinning of bodily experience for the cognitive experiences of memory and understanding. The spiritual significance of these collective rites thus could be described as individual and primarily subjective.

Durkheim's emphasis, however, is clearly on ritual's social significance. The anticipation of and participation in these yearly ceremonies heightens the individual's sense of membership in the group: "they serve to remake individuals and groups morally" (1965, p. 414). The element of ritual in a sense combines the social and cognitive aspects of Durkheim's thought; it connects the inward-looking act of remembering with outward physical performance.

The remaking of these bonds, year after year, accumulates in memory; older people with more experiences of rituals see them differently than the young. For all celebrants, rituals are links with the past, if only in that they are performed in consistent ways from one year to the next. But for the elderly, whose memory extends back through time to celebrations of the same ritual when they were children, the reenactment also renews links with family and friends who have died. Thus rituals ensure survival of the social group over generations, in that it is the group's

own story that is reenacted. The rhythm of anticipation of these holidays evidently benefits the survival of the individual as well.

Is the meaning of these multiple facets of religious involvement likely to be different for the elderly than for the young or middle-aged? Several suggestions have been made thus far that it is. Older people frequently retain roles of authority in their congregations after they have retired or disengaged from other public roles. The material support offered by other members of the congregation in times of crisis such as bereavement may be more frequently called upon by the elderly. The developmental need for psychological transcendence of the body could be facilitated by a heightened spiritual identity. Ritual observances, too, could be more important for the elderly; the repetition of these ceremonies are sedimented in memory, building up a sense of continuity over the lifetime in weekly and yearly cycles.

A fourth and final area of findings takes our work back to Durkheim. Our findings regarding the differential effects of public religiousness on both disability and the timing of mortality indicated that primarily Catholics and, to a certain extent, Jews benefit from their religious observance. These effects were minimal among black Protestants and nonexistent among white Protestants. This rank ordering is very similar to that given long ago by Durkheim (1951) and recently by Pescosolido and Georgianna (1989) in their studies of the risk of suicide. Durkheim's account of these differences focuses on "the existence of a certain number of beliefs and practices common to all the faithful, traditional and thus obligatory" (1951, p. 170). Pescosolido and Georgianna (1989) recast Durkheim's work with network theory, but like Durkheim found that Protestants both then and now killed themselves more often because their social organizations were more open to a spirit of free inquiry, a condition that results in fewer constraints in social ties, less ritualistic practices in worship, less authority of clergy and church dogma, in short, a less integrated society with an insufficiently intense collective life. Our work suggests (1) that the health benefits of Catholicism and Judaism extend beyond these well-known results regarding suicide and (2) that the key to understanding the varying intensities of these collective lives may be in understanding what is uniquely and particularly religious about them: the ritualistic enactment of their beliefs.

In sum, we see a powerful phenomenon at work here. In almost every way we examined the relationship, religious involvements on both private subjective and public behavioral levels were linked with the preservation of life and health, as Durkheim said they were. The ability to perform the fundamental activities of daily life, the prevention of feelings of emotional distress, and even the ability to control the timing of death

all appear to be benefited by one or another aspect of the religious experience of this group of elderly people. The findings for this sample should encourage more work in an area that has had surprisingly little attention from social scientists, given the availability of adequate empirical methods for analyzing these complex and subtle relationships and the centrality of the ideas to classical sociological theory.

REFERENCES

Berkman, Lisa F., and S Leonard Syme. 1979. "Social Networks, Host Resistance, and Mortality: A Nine-Year Follow-up Study of Alameda County Residents." *American Journal of Epidemiology* 109 186–204.

Berkman, Lisa F., et al 1986. "Depressive Symptoms in Relation to Physical Health and Functioning in the Elderly." *American Journal of Epidemiology* 124:372–88.

Cornoni-Huntley. Joan, et al. 1986 *Established Populations for Epidemiologic Studies of the Elderly: Resource Data Book* NIH Publication no. 86-2443. Rockville, Md.. National Institute on Aging

Durkheim, Emile. (1915) 1965 *The Elementary Forms of the Religious Life* New York: Free Press.

———. (1897) 1951. *Suicide.* New York. Free Press.

Higgins, Millicent W., and Jacob B. Keller. 1973. "Seven Measures of Ventilatory Lung Function Population Values and a Comparison of Their Ability to Discriminate between Persons with and without Chronic Respiratory Symptoms and Disease, Tecumseh, Michigan." *American Review of Respiratory Disease* 108·258–72.

Holt, Mary M. (1977) 1982 Revision. *SURREGR: Standard Errors of Regression Coefficients from Sample Survey Data.* Research Triangle Park, N.C.. SAS Research Triangle Institute.

House, James S., Cynthia Robbins, and Helen L Metzner. 1982. "The Association of Social Relationships and Activities with Mortality. Prospective Evidence from the Tecumseh Community Health Study." *American Journal of Epidemiology* 116:123–40.

House, James S , Karl R. Landis, and Debra Umberson. 1988. "Social Relationships and Health." *Science* 241.540–45.

Idler, Ellen L. 1987 "Religious Involvement and the Health of the Elderly. Some Hypotheses and an Initial Test." *Social Forces* 66:226–38.

Idler, Ellen L., Stanislav V. Kasl, and Jon Lemke. 1990. "Self-Evaluated Health and Mortality among the Elderly in New Haven, Connecticut, and Iowa and Washington Counties, Iowa, 1982–1986." *American Journal of Epidemiology* 131: 91–103

Information Please Almanac Atlas and Yearbook 1981. New York Simon & Schuster.

Jarvis, George K., and Herbert C. Northcott. 1987 "Religion and Differences in Morbidity and Mortality." *Social Science and Medicine* 25:813–24.

Katz, Sidney, et al. 1983 "Active Life Expectancy." *New England Journal of Medicine* 309.1218–24.

Larson, David B , et al 1986 "Systematic Analysis of Research on Religious Variables in Four Major Psychiatric Journals, 1978–1982." *American Journal of Psychiatry* 143:329–34

Lee, Eun Sul, Ronald N. Forthofer, and Ronald J. Lorimor. 1989 "Analyzing Complex Survey Data." Sage University Paper Series on Quantitative Applications in the Social Sciences, no. 07-071, Beverly Hills, Calif.

Levin, Jeffrey S., and Preston L. Schiller 1987. "Is There a Religious Factor in Health?" *Journal of Religion and Health* 26 9–36

McGuire, Meredith 1985 "Religion and Healing." In *The Sacred in a Secular Age: Toward Revision in the Scientific Study of Religion,* edited by P. Hammond. Berkeley and Los Angeles: University of California Press.

Mechanic, David. 1990. "Promoting Health." *Society* (January/February), pp. 16–22

Pargament, K. 1982. "The Interface among Religion, Religious Support Systems, and Mental Health." In *Community Support Systems and Mental Health,* edited by D. Biegel and A. Naparstek. New York: Springer.

Peck, Robert C. 1968. "Psychological Developments in the Second Half of Life." Pp. 88–92 in *Middle Age and Aging,* edited by Bernice L. Neugarten. Chicago. University of Chicago Press.

Pescosolido, Bernice A., and Sharon Georgianna. 1989. "Durkheim, Suicide, and Religion: Toward a Network Theory of Suicide." *American Sociological Review* 54.33–48.

Phillips, David. 1978. "Deathday and Birthday· An Unexpected Connection." Pp. 71–85 in *Statistics: A Guide to the Unknown,* 2d ed Edited by Judith Tanur et al. San Francisco. Holden-Day.

Phillips, David, and Kenneth A Feldman. 1973. "A Dip in Deaths before Ceremonial Occasions. Some New Relationships between Social Integration and Mortality." *American Sociological Review* 38.678–96.

Phillips, David, and Elliot W. King. 1988. "Death Takes a Holiday: Mortality Surrounding Major Social Occasions." *Lancet* September 24 (2): 728–32.

Phillips, David, and Daniel Smith. 1990. "Postponement of Death until Symbolically Meaningful Occasions " *Journal of the American Medical Association* 263:1947–51.

Rose, Geoffrey A 1962 "The Diagnosis of Ischaemic Heart Pain and Intermittent Claudication in Field Surveys." *Bulletin of the World Health Organization* 27:645–58.

Schoenbach, Victor J., Berton H. Kaplan, Lisa Fredman, and David G. Kleinbaum. 1986. "Social Ties and Mortality in Evans County, Georgia." *American Journal of Epidemiology* 123:577–91.

Siegel, Judith M., and David H. Kuykendall. 1990. "Loss, Widowhood, and Psychological Distress among the Elderly." *Journal of Consulting and Clinical Psychology* 58.519–24.

Shah, B. V , R. E. Folsom, F. E Harrell, and C. N. Dillard. 1984. "RTILOGIT. Procedure for Logistic Regression on Survey Data." Research Triangle Park, N.C . SAS Research Triangle Institute.

Taylor, Robert J., and Linda M. Chatters. 1986. "Patterns of Informal Support to Elderly Black Adults: Family, Friends, and Church Members." *Social Work* 31:432–38.

Umberson, Debra. 1987 "Family Status and Health Behaviors. Social Control as a Dimension of Social Integration " *Journal of Health and Social Behavior* 28.306–19.

Witter, Robert A., William A. Stock, Morris A Okun, and Marilyn J. Haring 1985. "Religion and Subjective Well-Being in Adulthood: A Quantitative Synthesis " *Review of Religious Research* 26:332–42

Zerubavel, Eviatar. 1981. *Hidden Rhythms: Schedules and Calendars in Social Life.* Chicago University of Chicago Press

———. 1982. "Easter and Passover. On Calendars and Group Identity." *American Sociological Review* 47:284–89.

Zuckerman, Diana M., Stanislav V. Kasl,·and Adrian M Ostfeld. "Psychosocial Predictors of Mortality among the Elderly Poor " *American Journal of Epidemiology* 119:410–23

The Political Empowerment and Health Status of African-Americans: Mapping a New Territory[1]

Thomas A. LaVeist
Johns Hopkins University

This article examines the impact of black political gains on black and white postneonatal mortality. Two dimensions of black political empowerment are examined, absolute political power and relative political power. The analysis examines all U.S. central cities with a population of at least 50,000 residents, 10% of whom are black. Absolute black political power did not influence postneonatal mortality for blacks or whites. However, there is a negative association between relative black political power and the black postneonatal mortality rate. Black political power had no significant effect on white postneonatal mortality. The implications of these findings for medical sociology theory are discussed.

INTRODUCTION

For decades, sociological research on health has employed an established set of social variables such as age, sex, and marital status and personal attributes that reflect attitudes and behavior. Political power, however, has not been among them. Recently, the idea that political and community empowerment leads to improved health status of the empowered group has received some attention from health social scientists. However, recent published discussions have mainly appeared in the form of speculative essays (McKnight 1985), letters to the editor (Braithwaite and Lythcott 1989), or as case studies with indeterminable generalizability (Brad-

bury 1987; Miller 1987). Social science has yet to attempt to assess systematically the hypothesis that political empowerment brings with it beneficial health consequences. It is also important that this hypothesis be tested within the calculus of socioenvironmental factors known to affect health status. This article attempts to assess this set of relationships empirically.

THEORETICAL FRAMEWORK· THE SOCIAL CAUSES OF DEATH

Since the early 20th century, researchers have observed a link between standard of living and morbidity and mortality (Ogburn and Thomas 1922; Tugan-Baranowsky as quoted in Thomas 1925; Newsholme 1910). In his seminal book, *Mirage of Health,* Dubose (1959) credited an increased general standard of living, economic development, and superior sanitation for reducing mortality rates. Support for this perspective has been wide-ranging (McKinlay, McKinlay, and Beaglehole 1988, 1989; Pampel and Pillai 1986; McKeown 1976; McKinlay and McKinlay 1977; Rydell 1976; Brenner 1973). Since Illich's (1975) elaboration, this theory has come to be the dominant medical sociology theory of mortality decline (McKinlay, McKinlay, and Beaglehole 1989; Conrad and Kern 1986, pp. 7–9).

The theory holds that social factors have been primarily responsible for reducing mortality rates. These factors include societal transitions in economic conditions, urbanization, improved sanitation, lower fertility, and improvement in the status of women. Modernization led to greater urbanization, which improved access to medical services and to quality housing (Rydell 1976). Higher incomes and an improved standard of living led to improved nutrition and a sanitary environment more conducive to good health among mothers and their infants (Dubose 1959). Modernization also led to fertility decline and thus to an older maternal age at birth (Knodel and Hermalin 1984). And a better educated population led to attitudinal and life-style changes that promote good health (Showstack, Budetti, and Minkler 1984).

Work supportive of this theory comes primarily from research on developing countries or analyses of European experiences during the 19th century (Berrebi and Silber 1981). However, analysis of industrialized nations provides evidence that the marginal return of economic development in reducing mortality may have reached the point of diminishing returns. Preston (1975) raised the possibility that beyond a threshold at which economic standing, sanitary conditions, nutrition, and so on are supportive of better health status, the return on increasing increments in the standard of living ceases to reduce mortality. In fact, the slow rate

of decline in infant mortality in the United States since the 1960s suggests some evidence in support of Preston.

Although the absolute risk of infant death in the United States is rather low, there are substantial differences between various class-race categories. For example, the infant mortality rate for African-Americans is roughly double that of whites. It has been empirically demonstrated that this race disparity is in large part indicative of social inequality (LaVeist 1990, 1989). It has also been hypothesized that this disparity is reflective of underlying political inequality (McKnight 1985). "Poverty and powerlessness create circumstances in people's lives that predispose them to the highest indexes of social dysfunction, the highest indexes of morbidity and mortality, the lowest access to primary care, and little or no access to primary preventive programs. Poverty of the spirit and of resources remains *the* antecedent risk factor of preventable disease" (Braithwaite and Lythcott 1989, p. 282).

The hypothesis that political empowerment will lead to improved health status takes on importance in light of the tremendous political gains African-Americans have made during the past two decades. The total number of African-American elected officials increased from 1,469 in 1970 to 6,424 in 1986. The most impressive gains have been among municipal officials whose numbers have increased from 623 in 1970 to 3,112 by 1986 (Joint Center for Political Studies 1986). If it is indeed true that race differences in mortality are manifestations of underlying powerlessness, it stands to reason that where African-Americans have gained political power black mortality rates should be lower.

DATA

In this article, I use U.S. central cities as the unit of analysis. The selection criteria included all U.S. central cities that, in 1980, had a population of at least 50,000, 10% of which was black. These criteria were selected to conform to previous usage in other studies (Jiobu 1972; Marston 1968; Marshall and Sinnott 1972; Dye and Renick 1981). This selection process resulted in a population size of 176 ($N = 176$) cities in 32 states.[2] Data

---

[2] Outlier analysis indicated that there were two cities that have extremely high postneonatal mortality rates yet have very low, low birth-weight rates In both cases the magnitude of these cities' deviation from other cities is such that they exerted great influence in the analysis and resulted in a nonintuitive inverse relationship between low birth-weight rate and postneonatal mortality rate. These cities were deleted from the analysis and a positive relationship resulted (although low birth weight was still not statistically significant in the black model). In light of this, I have excluded these two cities from the analysis. These outliers explain a nonintuitive finding reported in a related paper (LaVeist 1989), however, the conclusions in that paper are unaffected

on black city council representation come from the annual report published by the Joint Center for Political Studies (1986) and were supplemented by telephone calls to the various city councils. Black and white postneonatal mortality rates are derived from vital statistics reports (U.S. Department of Health and Human Services 1981, 1982, 1983, 1984, 1985). Poverty and education data come from census reports (U.S. Bureau of the Census 1980). Segregation indices are the index of dissimilarity (White 1986). And low birth weight and unwed birth data are unpublished data provided by the National Center for Health Statistics.

MEASUREMENT

In this study the postneonatal mortality rate is used as a general indicator of health status. Postneonatal mortality refers to deaths that occur between the second and twelfth months. Such deaths are generally thought to be primarily influenced by social factors such as poverty and environmental hazards (LaVeist 1989; Starfield 1985). Infant death rates are a well-established indicator of overall socioeconomic development, the availability and utilization of health services, the health status of women of childbearing age, and the quality of the social and physical environment (Morris 1979). Postneonatal mortality rates are selected instead of general infant mortality rates because postneonatal deaths usually occur after the infant has been brought home from the hospital and has been exposed to the mother's living conditions. In this sense, postneonatal mortality is a more sensitive indicator of general living conditions and is a broad index of health status resulting from socioenvironmental factors.

Postneonatal mortality rates are the total number of race-specific postneonatal deaths over the total number of race-specific births per 1,000 live births. All rates are based on births and deaths occurring during the five-year period 1981–85. Five-year rates were used to reduce the instability that may be caused by small numbers of postneonatal deaths in some of the smaller cities.

Black political power is measured in two ways in this analysis. The first formulation conceptualizes black political power as political representation of black political power relative to black voting potential (Browning, Marshall, and Tabb 1984). Relative black political power is measured as the proportion black on the city council (PBCC) divided by the proportion black in the voting-age population (PBVAP).[3] If PBCC is

[3] Preliminary analysis of these models used variations of the measurement of black political power These variations included a subtraction method used by MacManus (1978) and Taebel (1978). In this formulation, the percentage of blacks in a city's population was subtracted from the percentage of blacks on that community's city

less than PBVAP, then the resulting ratio score will be less than one and the African-American community of the city is considered politically underrepresented. If PBCC is greater than PBVAP, the city's black population is considered overrepresented, thus more politically empowered. A city in which PBVAP and PBCC are equal is proportionally represented. Such a city's black population is in political equilibrium and is, therefore, considered politically empowered.

The second measure of black political power is an indicator of absolute black political empowerment. This measure is the percentage of city council members who are black. This formulation refers to the level at which African-Americans are empowered to control the political and policy-making apparatus of the city. It is a plausible alternative hypothesis that the absolute percentage of black city council members would be a more potent predictor of health status than relative political power. It is plausible that the presence of a voting majority on the city council regardless of the makeup of the voting-age population would lead to changes in the allocation of funds or policy initiatives that benefit black constituents.

Several known correlates of infant and postneonatal mortality are included in the model, primarily as control variables.

1. The poverty rate is the proportion of a city's families living below the federally defined poverty limit. This variable is race specific.
2. The index of dissimilarity is a measure of racial residential segregation. The index ranges from 0 to 100 and measures the degree to which African-Americans and whites are residentially segregated within the city. It can be interpreted as the percentage of African-Americans that would have to relocate to another census tract so that every census tract would have a black/white proportion equal to the black/white proportion for the entire city (Taeuber and Taeuber 1965, Duncan and Duncan 1955, Deskins 1972) (For a full conceptual discussion of the index and other measures of segregation, see White [1986, 1983]).
3. The unwed birthrate is the number of infants born to unmarried women per 100 live births for the years 1981–85. This variable is race specific.
4. The low birth-weight rate is the race-specific number of low birth-weight births (less than 2,500 gm) per 100 live born infants in each city for the years 1983–85

---

council. A variation on the subtraction method used in the preliminary analysis was to use the percentage of African-Americans in the voting age population instead of that in the total population. Preliminary analysis also examined the regression method first reported by Engstrom and McDonald (1981) and replicated by Vedlitz and Johnson (1982). A dummy variable indicating whether the mayor was black was also specified in preliminary analysis. The various specifications of political power varied little in their effectiveness in predicting postneonatal mortality. The ratio method reported in this paper was selected as the measure of relative black political power because it is the best-established method in the literature.

5. Black and white educational levels are measured by the percentage of female high school graduates.
6. Geographic region of the country is based on the U.S. Postal Service region classifications (Northeast, North Central, South, and West) and is included in the analysis as a set of dummy variables interpretable in relation to the Northeast region.
7. Percentage black and log of city's population are also included in the analysis.

ANALYSIS

My analysis begins with an exploration of the univariate statistics for the variables included in the models. It is also important to assess the black/white differences among these variables. This analysis supports previous findings that, in these social indicators, African-Americans are at a general disadvantage compared with whites.

Table 1 displays means and standard deviations for the exogenous variables and the dependent variables. The table shows that the mean black postneonatal mortality rate among these cities is roughly double the white race. The mean black postneonatal mortality rate is 6.3 deaths per 1,000 live births. This is 1.8 times greater than the white rate, which is 3.5. The average city in this study has a black city council representation of 22.5%. The average black political power ratio is .82. Therefore, it can be said that although African-Americans have experienced substantial gains in political representation, they have yet to reach their political potential. The black population of the average city has reached only 82% of its potential political power.

Poverty is substantially more prevalent among African-Americans. The black family poverty rate in 1980 was 3.8 times greater than the white rate. The average city had a family poverty rate of 26.1 among blacks and 7.1 among whites. In the average city, 66.6% of white females and 53.7% of black females are high school graduates. The mean racial residential segregation score among these cities is 64.8.

The unwed birth rate is 3.8 times higher for the black population of these cities than for the white population. The average five-year unwed birthrate for blacks was 58.0 out-of-wedlock births per 100 live births, for whites it was 15.1. And the black and white low birth-weight rates are 16.2 and 6.8, respectively, per 100 live births.

BLACK POLITICAL POWER AND POSTNEONATAL MORTALITY

The analysis now turns to a multivariate test of the relationships hypothesized above. Table 2 examines the effect of black political power on black postneonatal mortality. The model is fitted using race-specific indepen-

TABLE 1

MEAN AND STANDARD DEVIATIONS FOR DEPENDENT AND CONTROL VARIABLES

Variable	Mean	SD
Postneonatal deaths per 1,000 live births:*		
Black	6 3	2 2
White	3.5	1.1
Ratio	1 8	
% black in city council †	22.5	11 9
Black political power ratio †	82	.1
Family poverty rate ‡		
Black	26 1	6.8
White	7.1	2.7
Ratio	3.8	
% female high school graduates ‡		
Black	53 7	9.6
White	66.6	10 2
Ratio	8	
Residential segregation index ‡	64.8	13 4
Unwed births per 100 live births *		
Black	58.0	11.1
White	15 1	9 0
Ratio	3.8	
Low birth-weight births per 100 live births.§		
Black	16 2	8.3
White	6.8	3 8
Ratio	2 4	

* Aggregated for the years 1981–85
† 1985
‡ 1980.
§ Aggregated for the years 1982–85

dent variables. Black postneonatal mortality rate is modeled as a linear function of relative and absolute black political power, poverty, unwed birthrate, low birth-weight rate, education, and residential segregation. Natural log of population, percentage black, and region are also controlled in the analysis. The table displays the results of this analysis.

The analysis in table 2 confirms the hypothesized inverse effect of relative black political power on black postneonatal mortality. That is, a higher level of black political power relative to the black voting-age population is associated with lower black postneonatal mortality rates. Although relative black political power is associated with black postneonatal mortality, absolute black political power is not. There is an inverse, but nonsignificant, association between absolute black political power and the black postneonatal mortality rate.

The table further reveals that poverty has the predicted positive effect

TABLE 2

BLACK POSTNEONATAL MORTALITY REGRESSED ON BLACK POLITICAL POWER

Independent Variable	Relative Power	Absolute Power
Black political power	−.014 (.002)**	− 009 (.009)
Poverty	086 (03)***	086 (.03)***
Segregation	030 (01)**	027 (.02)*
Education	− 026 (03)	− 027 (026)
Unwed birthrate	.001 (00)	001 (.00)
Low birth weight	002 (00)	002 (00)
ln of population	189 (49)	153 (.50)
% black	214 (1 12)	421 (1 11)
West	− 036 (76)	− 093 (76)
North Central	1.166 (51)*	1 138 (.51)**
South	189 (49)	.168 (50)
Intercept	−1.529	−1 790
R^2 (adjusted)	11	10
F-ratio	2 78***	2 69***

NOTE.—Coefficients are unstandardized SEs are in parentheses T-ratio may be calculated by dividing the unstandardized (metric) coefficient by the SE

* $P \leq .10$
** $P \leq .05$
*** $P \leq 005$

on black postneonatal mortality rates. The effect of poverty is consistent in each model. Also, racial residential segregation has a significant positive effect on black postneonatal mortality. Higher levels of poverty and segregation lead to higher black postneonatal mortality rates. Finally, education, low birth weight, and unwed birthrate have coefficients in the expected directions.

Table 3 shows the results of the analysis for the white postneonatal mortality rate. The table shows that both relative and absolute black political power have a positive association with white postneonatal mortality. However, in neither case is black political power statistically significant. Thus it is reasonable to conclude that whatever gains African-Americans have received from political empowerment have not come at the expense of whites.

Table 3 also shows that unwed birthrate and low birth-weight rate have the expected positive association with white postneonatal mortality. Cities with higher white unwed birthrates and low birth-weight rates have higher white postneonatal mortality rates. Also, the analysis supports the well-established link between poverty and postneonatal mortality. Each of these relationships displays consistency across both models.

One other compelling finding in the analysis is the different ways in which racial residential segregation affect postneonatal mortality. Table 2

TABLE 3

WHITE POSTNEONATAL MORTALITY REGRESSED ON BLACK POLITICAL POWER

Independent Variable	Relative Power	Absolute Power
Black political power	001 (001)	.002 (004)
Poverty	149 (035)***	.145 (.035)***
Segregation	− 011 (007)*	− 01 (.007)*
Education	− 009 (.009)	−.007 (009)
Unwed birthrate	002 (.001)**	.002 (.001)**
Low birth weight	007 (005)*	007 (005)*
ln of population	109 (219)	124 (219)
% black	211 (.562)	184 (555)
West	1 26 (.324)***	1 269 (324)***
North Central	1 007 (255)***	1 062 (.254)***
South	828 (.279)**	828 (279)***
Intercept	2 474	2 387
R^2 (adjusted)	26	26
F-ratio	6 44***	6 44***

NOTE —Coefficients are unstandardized SEs are in parentheses T-ratio may be calculated by dividing the unstandardized (metric) coefficient by the SE.
* $P \leq$ 10
** $P \leq$ 05
*** $P \leq$ 005.

shows a positive association between segregation and black postneonatal mortality. And, in table 3, segregation has a negative association with white postneonatal mortality. This finding is discussed elsewhere (La-Veist 1989). However, in short, this finding is evidence that differences in the socioenvironmental conditions in which African-Americans and whites live are, in part, responsible for the 2 : 1 ratio in black and white postneonatal mortality.

It is important to comment on the mechanisms by which political power may be producing lower black postneonatal mortality. Perhaps the most obvious explanation is that black political officials distribute municipal resources in such a way as to benefit black communities. Some support for this proposition is found in the literature. Karnig and Welch (1980) found that municipal budgetary allocations in cities with higher levels of black political power differed from cities with less black political empowerment. The high black empowerment cities had higher expenditures for sewers, social welfare, health, hospitals, and protective services. Thus, it could be said that black politicians used budgetary allocations to attempt to improve the general quality of life of their black constituents. Dubose (1959) demonstrated how improvements in sanitation, economic development, and the general standard of living led to lower mortality rates. Therefore, government spending in areas like sewers, social wel-

TABLE 4

Unstandardized Coefficients for per Capita Spending for Different
Municipal Services Regressed on Black Political Power and Control
Variables[a]

Variable	Relative	Absolute
Health . . .. -	014	013
Police	011	.014
Fire	046*	.040*
Streets	030*	.021
Sewers	.088**	.083**

[a] Regression analysis includes residential segregation, poverty rate, and region Coefficients refer to
separate analysis run on each expenditure
* P ≤ 10
** P ≤ 05

fare, and other areas to enhance quality of life may have positive conse-
quences for health status.

Table 4 displays the results of the analysis of this proposition. Per
capita municipal spending for health and hospitals, police, fire, streets,
and sewers were each regressed on black political power, residential seg-
regation, poverty, and region. This analysis was conducted for the mea-
sures of relative and absolute political power in separate models. Each
entry in the table represents a separate regression analysis. Each measure
of municipal spending was included as a dependent variable with the
political power variables included as independent variables. The table
shows unstandardized regression coefficients for the variables of interest
for each of these models.

Table 4 indicates that, overall, there is some evidence that black
elected officials allocate municipal resources in different ways than their
nonblack counterparts. For the measure of relative black political power,
the analysis shows that there were greater municipal resources allocated
to fire, streets, and sewers. The analysis for absolute black political power
resulted in statistically significant findings for two of the three areas of
funding that are examined.

The analysis in table 5 shows, however, that these municipal alloca-
tions do not translate into lower postneonatal mortality rates. Table 5
shows the unstandardized regression coefficients for black and white
postneonatal mortality regressed on per capita municipal spending for
health, police, fire, streets, and sewers. Each model included black politi-
cal power, residential segregation, poverty, and region. In only one model
does municipal spending significantly affect black postneonatal mortality,
and the white postneonatal mortality rate is affected by three sources of
spending.

TABLE 5

UNSTANDARDIZED COEFFICIENTS FOR BLACK AND WHITE POSTNEONATAL MORTALITY
REGRESSED ON PER CAPITA SPENDING FOR DIFFERENT MUNICIPAL SERVICES AND
CONTROL VARIABLES[a]

Variable	Black Postneonatal Mortality	White Postneonatal Mortality
Health	−.003	− 016*
Police	− 001	− 003
Fire	−.014	− 009**
Streets	−.024**	− 011*
Sewers	−.002	− 001

[a] Regression analysis includes black political power, residential segregation, poverty rate, and region
Coefficients refer to separate analysis run on each expenditure
* $P \leq .10$
** $P \leq 05$

DISCUSSION

It has been established that the degree to which African-Americans are politically empowered relative to their potential political power is a more potent predictor of mortality than absolute political power. Also, the analysis has failed to establish a direct link from black political power to the distribution of resources to black and white mortality. Thus, the mechanisms through which black political power affects black health are more subtle and complex than the simple direct redistribution of funds. How can these findings be reconciled?

The analysis in this study and previous research lead to the theory that political power, resource distribution, and health status are manifest indicators of a higher order latent factor that may be labeled "community organization." The black political power/postneonatal mortality relationship is caused by a common underlying factor. Figure 1 schematically represents this theory. Amos Hawley's (1963) work on community organization informs us that "power is an attribute of a social system rather than of an individual." Thus a collective of politicized individuals can influence political decisions. Crain and Rosenthal (1967) demonstrated this by showing that more highly politicized communities were able to sway local political decision making in school desegregation and water fluoridation. And Clark (1968) demonstrated that politicized communities influenced budgetary expenditures for urban renewal.

Bullard (1983) showed similar findings regarding location of solid-waste disposal sites. Bullard's analysis revealed that 11 of 13 solid-waste sites in Houston were located in communities that had a black majority. The study concluded that community organization "force[d] the adoption
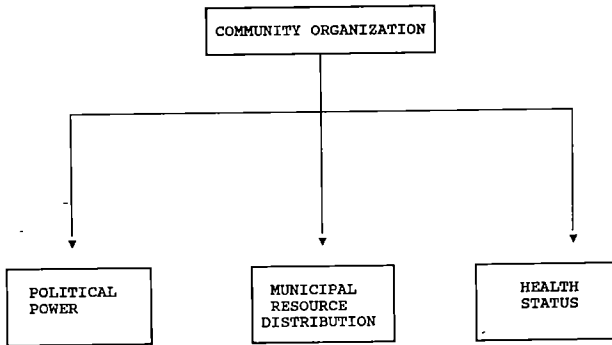
FIG. 1.—Theoretical model of black political empowerment

of alternative methods of waste disposal." If this is the case, one might expect an empowered black community to be a less toxic environment compared with a less empowered black community. And previous research has specified a link between environmental toxins and mortality and morbidity (Joyce, Grossman, and Goldman 1989; Hodgson 1970; Lave and Seskin 1970; Schwirian and Lacreca 1971).

Other researchers have found a greater inclination for community-level political participation among African-Americans in cities where blacks are politically empowered (Bobo and Gilliam 1990). In such cities one finds an African-American community with a strong community infrastructure (e.g., churches, civic groups, social organizations, civil rights organizations, block associations, etc.). Communities such as these are better able to elect black politicians (Karnig 1979). The infrastructure that makes these communities better able to elect black politicians may also make them better able to take advantage of opportunities, such as block grants and other programs, that provide services to community members (Friedman 1973). The general quality of life in these communities may thus be enhanced, and the improvement of social conditions may lead to lower postneonatal mortality.

Moreover, since these resource-rich communities (relatively speaking) are better able to influence local political elections, a greater number of black elected officials are to be expected in these cities. Thus black community leaders have an increased likelihood of gaining access to political leaders, influencing local funding decisions, and improving city services to blacks. As A. Phillip Randolph declared in 1937, "Power is the product and flower of organization" (as quoted in Franklin 1984).

CONCLUSION

The analysis in this paper has revealed a cross-sectional association between black political power and black postneonatal mortality. Of course,

any interpretation that includes causal assertions is premature on the basis of such evidence. However, this finding does provide a starting point from which to explore the role of political factors in mortality research and medical sociology more broadly. These findings also hold promise in that they are consistent with previous studies that have shown that black political power has led to improvements in African-American quality of life in other spheres of living (e.g., Dye and Renick [1981] and Eisinger [1982] on black employment; Robinson, England, and Meier [1985] on appointments to local school boards). Moreover, the suggestion that political power leads to improvements in health status is particularly noteworthy when considered within the context of medical sociological theory of causes of mortality and morbidity.

The notion that socioenvironmental factors are the major contributors to negative health status has become the predominant view among medical sociologists (McKinlay et al. 1988, 1989; Conrad and Kern 1986, pp. 7–9). However, Preston (1975) has shown that these factors may be of lesser importance in late 20th-century highly industrialized societies. It is, therefore, important to explore ways in which medical sociological theory may be expanded to accommodate these realities.

This analysis extends present theory by demonstrating that, not only do social factors such as poverty affect mortality, but the political organization of society also has important consequences in determining who lives, who dies, and when. Race disparities in mortality are merely outward manifestations of underlying race differentials in political and social power.

In this paper I have attempted to distinguish a point at which political sociology and medical sociology merge. I have attempted to identify political power as new territory for medical sociology in the 1990s. This variable is potentially important as a contextual variable (Bobo and Gilliam 1990), as an endogenous variable (LaVeist 1988), or as an exogenous variable (as in this article). Also, the relevance of political power as a concept in health research is not limited to the study of African-Americans. Political empowerment is relevant in the study of other ethnic groups, gender, class, or in the analysis of any situation in which status and power differentials exist. It is a variable and concept around which important research remains to be done.

REFERENCES

Berrebi, Z. M., and Jacques Silber. 1981. "Health and Development. Socio-economic Determinants of Mortality Structure." *Social Science and Medicine* 15.31–39
Bobo, Lawrence, and Franklin D Gilliam. 1990. "Race, Sociopolitical Participation and Black Empowerment." *American Political Science Review* 84:377–93.

Bradbury, Ronald C 1987 "A Community Approach to Health Care Competition "
   Inquiry 24 253–65.
Braithwaite, Ronald L., and Ngina Lythcott. 1989. "Community Empowerment as
   a Strategy for Health Promotion for Black and Other Minority Populations." *Jour-
   nal of the American Medical Association* 261 282–83.
Brenner, Harvey. 1973 "Fetal, Infant and Maternal Mortality during Periods of
   Economic Instability." *International Journal of Health Services* 3.145–59.
Browning, Rufus P., Dale R. Marshall, and David H. Tabb. 1984. *Protest Is Not
   Enough*. Berkeley and Los Angeles· University of California Press.
Bullard, Robert D 1983. "Solid Waste Sites and the Black Houston Community."
   Sociological Inquiry 53 273–88.
Clark, Terry N 1968 "Community Structure, Decision-Making, Budget Expendi-
   tures, and Urban Renewal in 51 American Communities." *American Sociological
   Review* 33.576–93.
Conrad, Peter, and Rochelle Kern 1986. *The Sociology of Health and Illness. Critical
   Perspectives*. New York: St Martin's
Crain, Robert L., and Donald B. Rosenthal 1967. "Community Status as a Dimen-
   sion of Local Decision-Making." *American Sociological Review* 32:970–84
Deskins, Donald R 1972 *Residential Mobility of Negroes in Detroit, 1837–1965*.
   Michigan Geographical Publication no. 5. University of Michigan, Ann Arbor.
Dubose, René. 1959. *Mirage of Health*. New York: Harper & Row.
Duncan, Otis D., and Beverly Duncan 1955. "A Methodological Analysis of Segrega-
   tion Indexes." *American Sociological Review* 20:210–17.
Dye, Thomas R., and James Renick 1981. "Political Power and City Jobs. Determi-
   nants of Minority Employment." *Social Science Quarterly* 62:475–86
Eisinger, Peter K. 1982. "Black Employment in Municipal Jobs. The Impact of Black
   Political Power." *American Political Science Review* 76 380–92.
Engstrom, R. L., and M. D. McDonald 1981 "The Election of Blacks to City
   Councils: Clarifying the Impact of Electoral Arrangements on the Seat/Population
   Relationship " *American Political Science Review* 75 344–54.
Franklin, V. P 1984 *Black Self-determination: A Cultural History of the Faith of
   the Fathers*. Westport, Conn.. Lawrence Hill
Friedman, Judith. 1973. "Structural Constraints on Community Action: The Case of
   Infant Mortality Rates " *Social Problems* 21:230–45.
Hawley, Amos H. 1963. "Community Power and Urban Renewal Success." *Ameri-
   can Journal of Sociology* 68.422–31
Hodgson, T 1970 "Short-term Effects of Air Pollution on Mortality in New York
   City." *Environmental Science Technology* 4·589–602
Illich, Ivan. 1975. *Medical Nemesis: The Exploration of Health* London. Chalder &
   Boyars.
Jiobu, Robert M. 1972. "Urban Determinants of Racial Differentiation in Infant
   Mortality." *Demography* 9 603–15
Joint Center for Political Studies. 1986. *Black Elected Officials: A National Roster*
   New York: UNIPUB.
Joyce, Theodore J., Michael Grossman, and Fred Goldman. 1989. "An Assessment
   of the Benefits of Air Pollution Control. The Case of Infant Health." *Journal of
   Urban Economics* 25.32–51.
Karnig, Albert K. 1976. "Black Representation on City Councils." *Urban Affairs
   Quarterly* 12.243–50
———. 1979. "Black Resources and City Council Representation." *Journal of Poli-
   tics* 41.134–49.
Karnig, Albert, and Susan Welch 1980. *Black Representation and Urban Policy*
   Chicago. University of Chicago Press.
Knodel, John, and A Hermalin. 1984 "Effects of Birth Rank, Maternal Age, Birth

Interval, and Sibship Size on Infant and Child Mortality. Evidence From 18th and 19th Century Reproductive Histories." *American Journal of Public Health* 74:1098–1106

Lave, L. B., and E. P. Seskin. 1970. "Air Pollution and Human Health." *Science* 169.1003–8

LaVeist, Thomas A. 1988. "An Ecological Analysis of the Effects of Political Inputs on Health Outcomes: The Case of the Black-White Infant Mortality Disparity." Ph.D. dissertation. University of Michigan, Department of Sociology.

———. 1989 "Linking Residential Segregation with the Infant Mortality Race Disparity in U S. Cities " *Sociology and Social Research* 73:90–94.

———. 1990 "Simulating the Effects of Poverty on the Race Disparity in Postneonatal Mortality." *Journal of Public Health Policy* 11:463–73.

MacManus, Susan A 1978. "City Council Election Procedures and Minority Representation: Are They Related?" *Social Science Quarterly* 59 153–61.

Marshall, Harvey, and James Sinnott. 1972. "Urban Structure and the Black/White Fertility Differential Examination of the 'Assimilationist' Model." *Social Science Quarterly* 52:588–601.

Marston, Wilfred. 1968. "Socioeconomic Differentiation within Negro Areas of American Cities." *Social Forces* 48:165–75.

McKeown, Thomas. 1976. *The Modern Rise of Population* London. Edward Arnold.

McKinlay, J., and S. McKinlay. 1977. "The Questionable Contribution of Medical Measures to the Decline of Mortality in the United States in the Twentieth Century." *Milbank Memorial Fund Quarterly* 55:405–28

McKinlay, John, Sonja McKinlay, and Robert Beaglehole. 1988 "Trends in Death and Disease and the Contribution of Medical Measures." In *Handbook of Medical Sociology*, 4th ed. Edited by H. E. Freeman and S. Levine. Englewood Cliffs, N.J.: Prentice Hall.

———. 1989. "A Review of the Evidence Concerning the Impact of Medical Measures on Recent Mortality and Morbidity in the United States." *International Journal of Health Services* 19:181–207.

McKnight, John L. 1985. "Health and Empowerment." *Canadian Journal of Public Health* 76, suppl.: S37–S38.

Miller, Irwin 1987. "Interpreneurship· A Community Coalition Approach to Health Care Reform." *Inquiry* 24:266–75.

Morris, Morris D. 1979. *Measuring the Condition of the World's Poor: The Physical Quality of Life Index*. New York: Pergamon.

Newsholme, A. 1910. *39th Annual Report of the Local Government Board*. Cd. 5312.

Ogburn, W., and D. Thomas. 1922. "The Influence of the Business Cycle on Certain Social Conditions." *Journal of the American Statistical Association* 18:324–40.

Pampel, F. C., and V. K. Pillai. 1986. "Patterns and Determinants of Infant Mortality in Developed Nations, 1950–1975 " *Demography* 23:525–42.

Preston, Samuel. 1975. "The Changing Relation between Mortality and Level of Economic Development." *Population Studies* 29:231–48.

Robinson, Theodore P., Robert E. England, and Kenneth J. Meier. 1985. "Black Resources and Black School Board Representation: Does Political Structure Matter?" *Social Science Quarterly* 66:976–82.

Rydell, L. H. 1976 "Trends in Infant Mortality Medical Advances or Socio-economic Changes." *Acta Sociologica* 19:147–68

Schwirian, Kent, and Anthony Lacreca. 1971. "An Ecological Analysis of Urban Mortality Rates." *Social Science Quarterly* 52:574–87.

Showstack, Jonathan A., Peter P. Budetti, and Donald Minkler. 1984. "Factors Associated with Birthweight: An Exploration of the Roles of Prenatal Care and Length of Gestation." *American Journal of Public Health* 74:1003–8.

Starfield, Barbara. 1985. "Postneonatal Mortality " *Annual Review of Public Health* 6:21–40.

Taebel, Delbert. 1978. "Minority Representation on City Councils: The Impact of Structure on Blacks and Hispanics." *Social Science Quarterly* 59.142–52.

Taeuber, Karl, and Alma Taeuber. 1965. *Negroes in Cities: Residential Segregation and Neighborhood Change*. Chicago Aldine.

Thomas, Dorothy S. 1925. Pp. 30–34 in *Social Aspects of the Business Cycle*. New York: Routledge.

U.S Bureau of the Census. 1980. *Census of the Population: Characteristics of the Population,* pt. C, tables 126, 131, 132, 137.

U.S. Department of Health and Human Services, National Center for Health Statistics. Various years. *Vital Statistics of the United States.*

Vedlitz, Arnold, and Charles A. Johnson. 1982 "Community Racial Segregation, Electoral Structure and Minority Representation." *Social Science Quarterly* 63.729–36

White, Michael J. 1983. "The Measurement of Spatial Segregation." *American Journal of Sociology* 88.1008–18.

———. 1986. "Segregation and Diversity Measures in Population Distribution." *Population Index* 52:198–221.

Beyond Rational Choice: The Social Dynamics of How People Seek Help[1]

Bernice A. Pescosolido
Indiana University

A classic problem common to sociology and other social sciences revolves around how people make decisions. Some recent approaches start with and revise an individually focused, rational action framework. While this orientation to building transdisciplinary, multilevel models provides many insights, it fails to capture essential features of social life The social organization strategy (SOS) framework presented in this article offers a complementary approach to social action in general and decision making in particular. This orientation, a network and event-centered counterpart to rational choice, rests on fundamental principles that distinguish the discipline of sociology: social interaction is the basis of social life, and social networks provide the mechanism (interaction) through which individuals learn about, come to understand, and attempt to handle difficulties. This approach shifts the focus from individual "choice" to socially constructed patterns of decisions, including consultation with others. Utilization research in medical sociology serves as a case for reviewing theoretical approaches to decision making and provides the background necessary to a theoretical exposition of the SOS approach using data from the National Survey of Access to Medical Care (1975–76). The results support the utility of pursuing the SOS framework.

INTRODUCTION

By the early 1980s, many social scientists had come to recognize that the complexity of social life needs to be matched by a companionate complexity in the theories we construct, the data collection efforts we mount, and

the analytic tools we bring to bear on the interpretation of data. For Alexander (1990), this phase of postwar sociology stresses synthesis and reintegration, with some theorists focusing on integrating the now-divergent streams in sociology (e.g., Collins's [1981] microinteraction chains, Stryker's [1980] structural symbolic interaction, Maines's [1982] mesostructure, Cicourel's [1981] negotiated order perspective, Giddens's [1984, 1989] structuration theory). For Coleman (1990, p. 664), it demands a "new social science" which "must cross the traditional boundaries of disciplines within which knowledge is ordered," with other theorists focusing on the building of cross-disciplinary bridges (see also Etzioni's [1988] socioeconomics and Ostrom's [1989] public choice). This movement toward a synthetic perspective by various authors in various ways heralds, perhaps, the beginning of a mature social science.

Much of the boundary-maintenance rhetoric within sociology and across social science disciplines is being replaced by attempts to build a transdisciplinary, multilevel framework based on a contextualized view of rational actors engaged in purposive action. Some argue that the growing currency of this view represents a "paradigmatic struggle," a "new methodenstreit," or an "imperialist intrusion" within sociology (Etzioni 1988, p. ix; Swedberg 1989; Collins 1986, respectively). Their concern reflects, at least in part, a failure by sociologists to offer "a widely shared and powerful conception" of action which, for Mechanic (1990, p. 93), underlies the current dominance of rational-choice-based approaches in areas like health care.[2] The critical issues here lie not in the stale debate over the viability of rational choice logic. They center on whether to

---

[2] The lack of development of a unified perspective from sociology reflects a number of developments in the discipline's history Unlike economics where microeconomists virtually exiled economic historians in the late 1800s (Swedberg 1989) or anthropology where individualistic models were introduced relatively recently (Izmirlian 1971, Schneider 1974), sociology has always been more eclectic, tolerant of diversity, and, as a result, more bifurcated. The debate between structure and agency has been especially heated and enduring, producing distinct schools and subfields where little cross-fertilization has occurred (Boudon 1988, Stryker 1987). Moreover, survey methodology has led to an emphasis on individuals' social characteristics to the exclusion of explicating underlying mechanisms or constructing measurements for larger structures (Abbott 1988; Alba 1981, Coleman 1986, Collins 1986)

accept or reject the idea that a synthesis based on revised rational choice theory either (1) represents the only way to meld the seemingly incongruous insights of the social sciences in one overriding perspective of action or (2) fully exploits the most basic sociological insights in building bridges across disciplines and levels of analysis. I argue that the rational choice explanation, with economic psychology as the fundamental microdynamic, presents but one useful way of exploring social action and denies what is the most likely contribution of sociology to understanding decision making.

To advance this argument requires the development of a different but complementary orienting framework that starts with and makes problematic those very things that rational choice theorists have ignored or brought in to overlay the psychological and economic bases of their perspective. Despite all our intradisciplinary disputes on theoretical position and methodological approaches, sociologists do appear to agree on the primacy of social interaction in forming the very essence of social life, that is, of society and individuals, and on social structure as defining the bounds of the possible. The social organization strategy (SOS) framework proposed here is social network centered and event centered. It begins with these sociological premises about human nature and social life stated above and overlays them with notions of utility maximization, purposive action, and bounded rationality derived from psychological and economic theory. The primacy of social interaction in decision making and the systematic patterning of interaction networks form the proper analytic focus for study, allowing direct consideration of agency (albeit in revised fashion as the agent in interaction) and allowing notions of social structure and context to take on concrete meaning.

This article offers a first step toward the development of this sociological counterpart to the recently advanced transdisciplinary multilevel orientation to action that is based on a revised rational choice. This first step builds on the ideas of others, from network theorists to symbolic interactionists (e.g., Heise 1987; Laumann and Knoke 1987; Simmel 1955; Stryker 1980; White and Eccles 1986) and requires two critical tasks. First, it involves consideration of efforts within sociology to link subfields and levels of analysis—taking stock, broadly but not exhaustively, of our conceptual tools. Second, it requires a consideration of the insights of other social science disciplines including anthropology and history as well as economics and psychology.

Seeing sociological insights merely as corrective relegates the discipline to a minor status and supports pessimistic predictions of its demise (see Wolfe 1989). Building the social organization strategy (SOS) framework does not require rejecting or denying the utility of a rational-action-based

synthesis. The latter remains a particular and useful way of slicing through a problem (Gray 1987, p. 47). Its sociologically based counterpart provides another—recasting questions; calling for different methodological approaches; and providing new insights to social phenomena. The two strategies together might produce a more comprehensive understanding of social action, each focused on the same general phenomena with lenses aimed at different angles. The two orientations also share the same goal: to get beyond traditional bifurcations within and across social science disciplines and develop a more encompassing perspective. Where they differ is in the starting point for action and in how to build a foundation for a transdisciplinary, multilevel framework. The synthesis based on sociology does not ignore individuals, their self-interest, their purposive action, or their rationality; it simply assigns them a different priority in action.

The issue of lay decision making for medical care provides an arena in which the utility of both approaches can be seen. This substantive focus offers two major benefits. First, the movement toward a counterpart orientation has been a conclusion reached more directly in medical sociology, in part because of a strong, unwavering multidisciplinary tradition and also in part because a number of shifts in the larger sociomedical profile of health, illness, and healing highlight the problems with traditional approaches. The wide range of cross-national studies of medical care choice using "thick description" (Janzen 1978), a "career" approach (e.g., Aday, Anderson, and Fleming 1980; Clausen and Yarrow 1955), and a "lay referral system" framework (e.g., Freidson 1970a; Furstenberg and Davis 1984; Horwitz 1977; Salloway and Dillon 1973) as well as the shift to greater societal burden of chronic illness generate an uneasiness with an individual, rationally focused perspective and provide the basis for a counterpart. Second, decision making for medical care offers a crucial arena for evaluating frameworks about decisions that is less biased in favor of rational choice than the economic arena where its relevance can be assumed to be more pronounced.

THEORETICAL BACKGROUND: SOCIAL SCIENCE MODELS OF ACTION AND DECISION MAKING

Reflecting the movement toward synthesis outlined in the Introduction, those adopting a rational choice perspective have been led to reconsider the effect of more social and dynamic factors in decision making. For example, the newer "public choice" or "political economy" school in political science (not to be confused with the neo-Marxist political economy approach in sociology; see Mueller [1976]) has rediscovered social

norms, influence processes, and the situation or social context.[3] In this revised framework, decisions in the social world are "purposive" and best seen as being made by individual social actors (including organizations and nations) mulling over the costs and benefits of a particular action in situations with variable characteristics and under a social structure that offers constraints and opportunities (operationalized as additional utilities). According to Lindenberg (1985): "the theoretical starting point is the *individual* who, having preferences and being confronted with constraints, has to make *choices*" (p. 99; emphasis in original). Generally, approaches in this vein share a view of the individual as an egoistic, rational, utility maximizer, but they drop assumptions unrealistic for noneconomic phenomena (those that deal with uncertainty and incomplete knowledge or variability in situations), and they reject notions of omniscience, structural irrelevance, and total atomism (e.g., Coleman 1990, 1987; Elster 1978; Marsden 1981; Ostrom 1989). The new strategically rational actors (as Elster [1979] calls them) take account of the environment and others and are involved in dynamic processes such as learning and reflection (Friedman and Hechter 1988; Lindenberg 1985; Mueller 1976; Radnitzky and Bernholz 1987).

This approach based on rational and purposive action by individuals is compelling for two reasons. First, it draws on an elegant, parsimonious, and cleanly deductive approach to decision making that corresponds to the most sophisticated developments in empirical modeling and methodology (Coleman 1988). It is, in fact, unmatched in its theoretical development by competing formulations. Second, seen in the light of sociology of knowledge, it matches the worldview of most people living in advanced capitalist democracies, which includes the ultimate importance of agency (whatever structural constraints may exist); the conjoint nature of modern society, autonomy, and rational thinking; and the everyday notions of the inescapable need to balance pros and cons. As Coleman (1990) asserts, this approach represents the theory of action used implicitly by most social theorists and by most people in the "commonsense psychology" they use to understand what they and others do (p. 5). Unless individuals are seen as puppets of sociocultural locations, as posed in many early sociological and anthropological theories, ultimately individuals either act or they do not; ultimately, even if bounded in their rationality or structural or normative constraints, they weigh beliefs and perceptions about the consequences of their actions.

---

[3] These current efforts are, in part, a reaction to overspecialization as evidenced by a renewed interest in the broad-based orientation of early social theorists that had been downplayed with the development of distinct social science disciplines (e g., Elster 1975 on Leibniz; Wolfe 1989, Etzioni 1988; Stryker 1980 on Adam Smith)

With this new inclusive framework and revised definition of central concepts, what quarrel can sociologists have with a synthesis based on a rational choice framework? The answer rests on the relative place and significance of social life. Specifically, I argue that in using this orienting approach, we strip individuals from society in five related ways. First, while reducibility to individuals' mental calculus must be acknowledged at some level, the critical question is whether this calculus is sociologically meaningful. Is it the individual or the individual in interaction (and the structure of interactional events) that is the appropriate and most basic unit of analysis? (See Tilly 1984.) This cannot be accommodated in the rational choice framework. As Etzioni (1988) points out, the assumption of free-standing individuals as the decision-making unit in that perspective "is much more than a working hypothesis; it is an article of faith grounded in a deep commitment to the value of liberty" (p. 10).

Second, by bringing in the social merely as an overlay to individual mental events, we abandon sociological notions of the pivotal process of interaction as the mechanism through which social phenomena, including but not limited to decision making, occur. Interaction with others forms an essential element in the dynamics of decision-making processes. No matter how sophisticated, for example, the Prisoner's Dilemma game cannot simulate what happens in decision making, except in the most rare and bizarre cases (Wolfe 1989, p. 43). According to Granovetter (1985, p. 486), this general approach to social relations "has the paradoxical effect of preserving atomized decision making even when decisions are seen to involve more than one individual."

Third, by focusing on individuals and their purposive action, we remove, almost without recognition, the embeddedness of problems and their solutions in the social network including the resort to social network contacts as decisions or actions in themselves. Problems are often not simply individual matters. Instead, they frequently represent a shock to a network that calls forth the purpose of or the need for action. Actions of other people cannot be conceptualized only as exogenous factors in rational choice (i.e., conditioning preferences) or as another utility in the individual cost-benefit calculus or as being rationally pursued by actors in certain situations. In the revised rational-choice-based perspective, relations among persons facilitate action but are not, themselves, seen as action. The consultation of network ties, however, minimally implies action and represents a choice or solution in many cases. These contacts are part of a set of choices that are socially organized and need to be made problematic.

Fourth, the general focus in this framework lies often on *a* choice (on occasion, a series of discrete choices), plucking it from the larger event that necessitates action and that sets in motion a course of action (Fjell-

man 1976). As Tuma and Hannan (1984, pp. 5 ff.) suggest, microeconomic assumptions may have diverted attention from processes and change. The basic issue is not whether the new revised rational-choice-based framework is dynamic in theory, but whether it sufficiently reflects social dynamics. As Tilly (1984, p. 32) notes, even game theory will not replace the need to put relationships rather than individuals at the center of theories of action or decision making. Dynamic rational choice models force one action at a time and ignore how sequences of events are patterned, contingent, and emergent.

Fifth, the conceptualization of action as rational choice raises questions of the applicability of this framework as a general orientation to social action. Certain actions are targeted as *choices* and others are ignored as *habits*. As Camic (1986, p. 1075) points out, this is a direct function of adopting the rational actor paradigm since habits "are not something one is at all prompted to investigate, or even notice, when one assumes that action always takes the form of a reflective weighing." But to the extent that social action proceeds through cultural routines rather than individual cost-benefit calculation, as Corsaro and Rizzo (1988) among others contend, then much social action remains outside the purview of the framework. Etzioni (1988) and Stryker (1980) share the notion that the neoclassical view has experienced, in Veatch's (1973) terms, "a generalization of expertise." That is, ideas relevant to economic behavior have been converted into a general perspective that encompasses all action and, in the process, neglects the vast domain of actions that are not economic.

In sum, even when norms, networks, and situations are brought in as additional items on the individual's checklist, social forces remain either restricted to those perceived or acknowledged by the individual or to an enumeration of additional constraints. To adopt this basis for a synthetic model of the social sciences is to focus on an isolated choice or set of circumscribed choices that encompasses a restricted subset of action where individuals make cost-benefit calculations. The critical dynamic relationship among individuals and their networks and the larger structures that form from and shape them are downplayed, even dismissed.

THE SOCIAL ORGANIZATION STRATEGY FRAMEWORK. BASIC SKETCH OF A COUNTERPART

The SOS framework provides an explicitly dynamic, network-centered, event-based approach built on a sociological foundation. As the citations below to basic building blocks indicate below, it draws liberally and broadly from a number of sociological research traditions, incorporating and organizing many sociological insights and arguments. Its starting

point lies in the fundamental social nature of individuals and action, with social relationships rather than individual mental events at the center of social life (Tilly 1984). The "self" is a social product, defined and developed in social interaction; "society" is an ongoing process, both shaping and shaped by interaction in social networks; and "action" is composed of substantively organized episodes of interaction rather than discrete individual acts (Fine and Kleinman 1983; Granovetter 1985; Knorr-Cetina 1981; Simmel 1955; Stryker 1980). Beginning with the primacy of social life provides the SOS framework with four basic, related building blocks drawn from diverse research traditions in sociology.

First, the "elementary actor" (Coleman 1990, p. 503) becomes social and pragmatic (which is what much anthropology as well as sociology suggests)—not isolated and ever-consciously rational. This does not mean that individuals become social dopes or social dupes. People are not unconscious; they are knowledgable, skillful actors with a "practical consciousness" that allows them to both improvise and routinize (Giddens 1984; Haines 1988; Heise 1989b). Second, in linking the symbolic interactionist tradition with network theory, we find an alternative to the dominant theoretical emphasis on the individual. If people are social, the focus of analysis shifts from the individual to the individual in patterned interaction with others. Drawing from Simmel, Dewey, and the more recent life-course perspective, we can conceptualize social life as dynamic streams of action with social interaction instead of mental calculation as the mechanism through which it proceeds. Third, with individuals embedded in an ongoing relational dynamic, the proper unit of analysis is the social network—but not simply as a linking mechanism between actor and structure or as a way of synthesizing embeddedness into an actor-oriented framework.[4] Interactional events replace individuals; social networks replace social atoms (Tilly 1984, p. 27). Fourth, global macrovariables tapping context (e.g., time, place) represent different substantive and structural networks. As such, they are not seen as isolated from action or individuals but as fundamentally tied to them and to their histories in particular. By systematically addressing embeddedness with a network focus (White and Eccles 1986), we can deal concretely with

---

[4] This approach diverges from Etzioni's (1988, p. 188) socioeconomics where collectivities, like communities, are the decision-making units. A network approach acknowledges the multiplicity of social circles to which individuals belong and the often divergent social norms and loyalties that create tensions as well as moral obligations (Simmel 1955; Fischer 1982, Wellman and Wortley 1986). This view also can accommodate explanations of socialization and internalization that Coleman (1990, p. 932) finds difficult to bring into the framework he develops (see, e.g., Pescosolido [1986], for a dynamic, network theory of socialization, Macy [1990] on learning theory, rational choice, and social movements)

Giddens's (1984) duality of structure or Sewell's (1988) dialectic theory of action. Individuals and structure, as "abstractions from ongoing interactions," become inextricably intertwined and cannot be understood apart from each other or from the networks that shape them (Stryker 1980, p. 2, see also Heise 1989b; Collins 1986, Knorr-Cetina 1981).

With this basic foundation, the SOS framework brings a different but complementary perspective to the understanding of human behavior. The revised rational choice perspective focuses on the individual but tilts the analytic lens upward in an attempt to capture other larger conditions (including networks) that affect the individual unit of analysis (e.g., Coleman 1990). The SOS perspective takes different aim: it focuses directly at the middle level of networks and interaction with a wide-angle lens capable of capturing some of the intricacy of the outer edges of individuals and social structures.[5] There is a difference between seeing social norms and social networks as influences on decisions or individuals, as rational choice theory suggests, and seeing social interaction in bounded networks as the mechanism that underlies action. In the approach I develop here, a particular action, choice, or decision is embedded in a social process where the network interactions of individuals not only influence preference formation and define the situation but also drive the process of deciding whether something is wrong, whether anything can be done about it, what should be done, and how to evaluate the results. By doing so, network interactions produce systematic structures and contents (or cultures) and sometimes become crystallized into organizations and institutions that, in turn, affect social interactions.

Applying this overall framework to decision making requires an explicit understanding of (1) how concrete problems or issues are dealt with in the ongoing stream of social life; (2) the dynamics that operate in interaction; and (3) the roles played by structure and content in such an approach. Without the first, the SOS framework becomes a theory of society, basically compatible with Giddens's (1984) scheme but woefully underdeveloped in comparison. Without the second, it shares the fundamental flaw Coleman (1988) identifies in most sociological theory—there is no "engine of action," no understanding of what drives microprocesses. Without the third, it reduces structure to the aggregation of microprocesses accompanied by somewhat global contextual limiting factors such as time, space, and place (e.g., Collins 1981; see also Hilbert's revision [1990]). The building of these three components of a SOS frame-

---

[5] Other synthetic efforts that attempt to link the actor with the larger structure (even through a network mechanism) simultaneously employ two lenses, one focused from the bottom and the other from the top, and attempt to capture the interaction of the two (e.g., Burt's [1982] structural theory of action).

work draws from diverse but well-established research lines in sociology. My primary task here is to integrate their insights.

The SOS Framework and Decision Making

The SOS approach frames the process of decision making in terms of the episode rather than the choice. Conrad (1987, p. 8) provides the initial premise: "What is sociologically most interesting about uncertainty is how people manage it." The initial focus of decision making is the event that necessitates action; the primary frame for study is the entire episode that encapsulates the actions surrounding the event (see Laumann and Knoke [1987] for a similar view). Once the study of decision making is framed in this way, "strategies" of action in Simon's (1976) sense (i.e., the patterns, combinations, or sequences of choices or decisions over the course of the episode), and how they are socially organized become the central phenomena to be explained. As life-course theorists and social historians have suggested, unless we reconceptualize the phenomena away from "a" decision or even a set of decisions to more of an emergent "multi-phased decision process" (Elder 1978), it is unlikely that we can, in fact, capture some semblance of ongoing history in our models (Sewell 1988).

The SOS approach assumes, following Laumann and Knoke (1987), that networks are antecedent to an event and that these networks do have a structure and content to them, although it makes problematic how individuals, in response to a particular event, choose to activate particular sectors of the multiple networks in which they are embedded (Aronson 1970; Huckfeldt and Sprague 1986). The overall system of relations, with its a priori structure and content, provides the basis for issues, problems, or other events and calls for specific attention (whether action or inaction). This is key—in the SOS framework, decision making itself is a dynamic, interactive process fundamentally intertwined with the structured rhythms of social life. A decision-making process, like other processes (e.g., social control), is "mounted on the back of ongoing social processes" (White and Eccles 1986, p. 131). Events set into motion a specific process of coping with uncertainty—initiated either by the focal individual's facing a problem or by network members who perceive a shift in the ongoing rhythm of social life. In this sense, events can be conceptualized as network phenomena rather than as simply individual phenomena. They can be seen as a "shock" to a network, reverberating through it and altering the overall system of relations. Dealing with problems can reaffirm, create, or destroy networks and can shift the ongoing trajectory of social life, sometimes in profound ways. In the ultimate case, it can create a new structure of social relations, change the content

of networks (e.g., beliefs, values, priorities), and, as a result, influence the cultural orientations, not only of a focal person or of those who remain connected to the individual, but of the society at large (on the latter point, see Wolfe [1989, p. 216]).

To understand these complex processes and structural ramifications we must separate, even if only analytically, the larger dynamic from the particular stream related to the event. First, the stream of network interactions surrounding a particular event (the episode) are much like "careers" (e.g., the "illness career," the "occupational career") in the life-course perspective and provide conceptual descriptions, not explanations (Elder 1978; see also Abbott and Hrycak 1990; MacKinnon and Heise, in press). Second, the theoretical explanation of decision-making episodes focuses on the larger, overall pattern of structured interactions or networks in which careers operate. Here individuals recognize or fail to recognize a problem, find the limits of social resources (e.g., knowledge, beliefs, instrumental aid), and find a way to evaluate the outcomes of action.

Adopting this approach allows for the conceptualization of both *static* hypotheses that freeze the larger stream to examine the effect of the overall set or system of networks on the career stream or a particular embedded act or sequence of action and *dynamic* hypotheses that allow the examination of how change affects the course of decision making. In illness, for example, having access to a dense network with beliefs skeptical of the efficacy of modern medicine increases both the resort to alternative healers and the delay in seeking out physicians (Freidson 1970a). And, if the illness "damages" network ties over time, perhaps through stigma or the burden of care, any continued compliance with a treatment regime is affected by the new mix of network density and ideology (see Pescosolido [1991] for a theoretical medical care model with static and dynamic hypotheses, based on an SOS framework).

Focusing Downward: The Microprocesses of Social Action

The SOS framework rejects the sole reliance on rational, cost-benefit calculation as the "engine of action" but does not eliminate it from consideration. As symbolic interactionists, ethnomethodologists, and cultural anthropologists have documented, people do not have to solve each problem anew or even understand the logic of old solutions; much human behavior is habitual, predictable, expected, taken-for-granted, and recurrent (Collins 1981; Corsaro and Rizzo 1988; Heise 1989a). Cultural routines, which form the basis of much day-to-day action, are largely acquired through association, "produced" through interaction, and dependent to a large extent on affective reactions (Corsaro 1990; Heise

1979; Kemper and Collins 1990; Stryker 1980). Affect underlies "acquired instinct" allowing people to sense what they should do without necessarily knowing why (Heise 1989b), to tap an "embodied history" (Bourdieu 1981) where much of the "cultural heritage of reasoned action" is stored (Heise 1987, p. 15), and to continually build for themselves a tacit dictionary, a font of experience and information, which makes the unfamiliar familiar. However, when cultural routines do not produce effective solutions to problems, individuals often become consciously aware of the need to think through situations. In short, much recent research suggests a multifaceted base of action—affect and rationality, emotion and cognition. In the SOS framework, action is seen as proceeding on the basis of both, each intrinsically social rather than individual (Collins 1981; Etzioni 1988, pp. 90 ff.). People do not necessarily engage in cost-benefit calculations when seeking information from others. Effective network ties are often neither costly nor of decreasing efficiency over time as Coleman (1990, p. 310) suggests, instead they are built into the routine of daily life (see e.g., Granovetter 1973).[6]

Focusing Upward: Structures and Decision Making

To argue that decisions are made through some mix of affect and cognition in interaction is not to suggest that the result is random, totally emergent, or idiosyncratic (Heise 1989b; Stryker 1980). The SOS framework makes it possible to link actors to each other, to the larger social system, and to such abstract entities as the state, the economy, and the community, which some sociologists claim represent the stable and recurring operation of social networks (Alba 1981; Fischer 1982; Laumann and Knoke 1987; Tilly 1984). This connection between action, interaction, and structure through networks needs to be elaborated at four levels—the relationship between network structure and (1) affect and interaction, (2) social characteristics, (3) ideas of context, and (4) the nature of institutions and organizations.

In the first, affect control theory provides a way in the SOS framework to link the "internal processing that generates social behavior to the sociocultural system that makes social interactions coherent" (Smith-Lovin 1987, p. 174). It does so through role identities: roles are not simply sets of behavioral guides but real identity-confirming actions where small

---

[6] Even here, the two are not incompatible Affect, as Etzioni (1988, pp. 103 ff.) points out, does not subvert mental processes, twist reason, or play a disturbing role in cost-benefit calculations, it often provides the necessary base Heise's work (1989a) suggests that affect may directly relate to assessment of utility in rational choice and thereby adjust rational processes

amounts of affectively encoded information give regularity to social interaction and link it to social structure (Heise 1979). As Bourdieu (1981, p. 309) notes, we engage in suitable role behavior "without ever having to 'pretend.' " In interaction, people see themselves and others as occupants of identities that evoke sentiments about moral value, power, and social energy. Maintaining these conceptions contributes as much to our understanding of action as rational calculation does (Heise 1987, 1979).

On social characteristics, urban anthropologists and network theorists have long contended and documented that among the most striking new aspects of complex societies is the importance of social relationships that take place outside bounded groups and institutions (Aronson 1970; White, Boorman, and Brieger 1976; Simmel 1955). Relying on the "cultural background of the participant" to measure the "extent to which individuals share norms" (Ostrom 1989, p. 17) falls into the structuralist trap of a static and homogeneous view of culture. Social characteristics can be bad proxies for social networks except under conditions of network homogeneity (similar to what Coleman [1990] calls closure). However, at least at this point in the development of the SOS framework, they hold a provisional status. Stryker and Serpe (1983) note that most social interaction does, in fact, occur between and among persons who are members of some common group. Social locations, tapped by characteristics such as age, sex, race, geographic, or organizational placement, partially shape the opportunity for network contact (Furstenberg and Davis 1984; Huckfeldt 1983; Marsden 1987; Whitten 1970). As we understand more about the nature of networks, their formation, stability, and operation, we will be able to describe more concretely their grounding in and divergence from sociocultural characteristics.

Concerning context, even though Tilly (1984, p. 17) charges us to discard the idea of society as a thing apart, I argue that we do need to treat network structures and their cultural content as such at some point. One key reason comes from rational choice theory. As Ostrom (1989, p. 20) notes: "the larger the set of individuals using the same resources and the more diverse their strategies, the more difficult it is for anyone to gain an accurate perception of another's strategic behavior." It may be that society is not a reality, sui generis, but empirically the expanse of network ties is so great and multilayered that we need to treat it as if it were (see Coleman 1990; Granovetter 1985; Ordershook 1986; for the similar status of rationality in their perspective). People do not, nor could they, monitor the set of ties in a church or school, yet, these structures affect social action (Freeman and Romney 1987; Killworth and Bernard 1976). To see them simply as global context variables such as place and time (e.g., Cicourel 1981; Collins 1981; Giddens 1984, 1989) misses an opportunity for understanding their substantive meaning. Context is, in

essence, a shorthand in the SOS framework for the operation of different network structures and contents, as ethnomethodological "indifference" to structure attests (Hilbert 1990). For example, what is different about rural and urban contexts in Taiwan is not the structure or operation of migrants' networks regarding medical care choices, instead, it is geographic context that taps variable network content regarding proper or ideal medical care, guiding migrants who are embedded in similar network structures in different geographic spaces to make dissimilar choices for either traditional or modern healers (Pescosolido 1986, see also Kadushin 1983).

On organizations and institutions, sometimes network interactions become so routine and patterned that they crystallize and, in turn, affect social life directly. And, in the SOS framework, physical and temporal barriers are seen as constraining or facilitating social interaction. For example, formal organizations affect important communication networks by structuring different daily work activities in separated physical locations. As a result, certain employees come into more frequent contact with some people but not others (e.g., see Kanter 1977; see also Knoke 1990, chap. 4). In this way, physical space, from the home to the nation-state, affects the formation and operation of networks. In social interaction, people do in fact bang up against real physical and temporal boundaries; to conceptualize these simply as global contextual factors again misses the opportunity to see how social interaction and networks are linked both to history and to social change.

In sum, the SOS approach builds a different but complementary framework for action that begins with assumptions about the social nature of individuals and their actions. Individuals are neither puppets of some abstract structure nor calculating individualists; people both shape and are shaped by social networks. The SOS framework sees interaction networks as the proper arena of analysis, the episode as the focus, affect and rationality as driving action in tandem, and social structure as tied fundamentally to network interaction patterns with a sui generis character. Incorporating the centrality of social networks demystifies and gives specific meaning to the idea of structure as a reality, particularly to the more chimeric notions of global contexts such as time and space.

THE STUDY OF MEDICAL CARE CONTACTS: ILLUSTRATING PROBLEMS AND PROVIDING DIRECTION

The study of individuals' medical care decision making (or utilization) as part of the subfield of medical sociology, offers a useful case to highlight the benefits and challenges of the SOS framework. Medical sociology, in general, has been relatively unfettered by interdisciplinary squab-

bles and has borrowed liberally from multidisciplinary traditions. Still its theoretical development broadly mirrors that of social science as discussed in the introduction. Early sociological and anthropological frameworks of medical care choice, although they differentially emphasized structure and culture, respectively, cast lay decision making in the structural-normative terms that echoed the larger historical reaction to microeconomics. Education, social class, and rural-urban location, for example, were conceptualized as reflecting the strength of traditional, nonscientific belief systems that discouraged the use of modern medicine (e.g., Koos 1954; Rivers 1927; Saunders 1954). However, very early on, interdisciplinary teams (e.g., Falk, Klem, and Sinai 1933) employed multifactor models and mounted national-level surveys. And Parsons's (1951) conceptualization of illness problems as a "sick role" drew great interdisciplinary attention and generated hundreds of research articles delineating contingencies of choice he had ignored (see Twaddle and Hessler [1977] for a review).

By the mid-1960s, two influential formal theoretical models organizing these important contingencies had been developed, with wide currency in the sociomedical sciences. It is interesting although not causal, that each initially reflected the polarity of approaches within sociology itself. The Health Belief Model (HBM) (Rosenstock 1966) focused on the social psychology of decision making, primarily the role of motivations, beliefs, and perceptions on individuals' decisions to seek formal medical care. The Socio-Behavioral Model (SBM) (Andersen 1968) was more structurally oriented—focusing primarily on access to and "need" for care but incorporating "predisposing" attitudes, beliefs, and characteristics. Over time, each model drew insights from the other and moved toward synthesis (see, e.g., Eraker, Kirscht, and Becker's [1984] "third generation" HBM; see also Stoner [1985] for an overview).[7]

What is striking about this synthesis is the extent to which it begins with the economic psychology of a rational choice perspective and is overlaid with corrections from other social sciences. Decision making is seen in preset dualistic terms—on the selection of one alternative, a modern medical professional, from the universe of potential sources for care, conceptualizing use or nonuse (or, alternatively, the volume of use) as the dependent variable. Given the rise of a modern profession of medicine based explicitly on notions of the superiority of scientific forms of medical

[7] Not only have these models held widespread transdisciplinary currency, they have also had great impact By successfully constructing a social profile of users and nonusers, they have been critical in the development of social policy in medical care, particularly during the 1960s and 1970s (see Aday 1972; Maurana et al. [1981] for comprehensive reviews).

care over any and all preexisting competitors (e.g., Larson 1977), this conceptualization of utilization as simply the decision to use a modern scientific practitioner, service, or facility reflected a "medical model bias" (Gold 1977) in medical sociology and "medical ethnochauvinism" in medical anthropology (Unschuld 1981). Along with these biases came the acceptance of "medicine's view of the patient as a singular, insular individual" (Conrad 1987, p. 6), which reinforced the utility of both random samples in survey data collection and sociodemographic characteristics as explanatory correlates (Freidson 1970b; McKinlay 1972).[8]

The most important alternative view to the dominant one, the "illness career" perspective, was pursued primarily by medical sociologists and anthropologists in the qualitative tradition, many of whom defined their research squarely in the grounded theory tradition. The notion of an "illness career" (a sequence of actions related to the attempt to rectify a health problem) has early, albeit limited, roots in medical sociology and anthropology. In sociological studies of these "pathways" to care, clergymen, police, lawyers, as well as friends and relatives, have been documented as critical actors in the social process of seeking care (Clausen and Yarrow 1955; Freidson 1970a; Gurin, Veroff, and Feld 1960; Kadushin 1966; Roth 1963; Suchman 1964). In addition, medical anthropologists, exploring choices in Western and non-Western countries, have documented the use of "alternative healers" such as shamans, chiropractors, curanderos, homeopaths. family, and others (Press 1969; Romanucci-Ross 1977; Rubel 1966; Unschuld 1976). Most recently, self-care, nonprescription, and home remedies have been incorporated as important options by sociomedical researchers (Ailinger 1977; Levin, Katz, and Holst 1976).

All of these eclectic research lines on the help-seeking process have broadened our view of the range of medical care options and the dynamic nature of lay decision making. By the mid-1970s, the image of competition among practitioners of different traditions with individuals making rational decisions to use one or another tradition was being replaced by an image of "complementarity" and simultaneous or sequential use

---

[8] As well as speeding the development of a synthesized multidisciplinary, multilevel perspective based on rational choice, medical sociology's transdisciplinary view simultaneously facilitated a critique of this framework and the development of alternatives. These commentaries and studies do not share the integrated and cumulative history of utilization models drawn above. Medical sociologists, like those in the discipline generally, tended to draw dichotomies (particularly in qualitative/quantitative terms) and to argue over superior approaches. But taken together, a number of research lines provide a substantial set of building blocks for a counterpart offering an expanded conceptualization of the nature and range of choices, the importance of process, and the central role of structural embeddedness.

throughout "events," "episodes," or "therapy courses" (Fabrega 1978, 1970; Janzen 1978; Unschuld 1976; Young 1981). By the mid-1970s, even developers of dominant utilization models based on a rational choice framework were incorporating these ideas as a way of dealing with questions of continuity, compliance, and delay in seeking care not amenable to analysis under their perspective (e.g., Aday et al. 1980, pp. 199 ff.).

In addition, four other lines of inquiry have contributed to the building of a counterpart approach. First, in spotty and somewhat unconnected fashion, a small network tradition of utilization research developed. Freidson (1970a) presented a first coherent framework for understanding the effect of network structure and content on the type of healer likely to be sought and the delay in seeking help. Following this, a few studies have marked the network paths to get to mental health clinics (Horwitz 1977; Kadushin 1966), child guidance clinics (Raphael 1964) and maternity clinics (McKinlay 1973). Second, in mostly unrelated developments from epidemiology, social and cognitive psychology, and social psychiatry, ideas of social support were offered as important in the etiology, treatment, and outcome of health and illness (see, e.g., Cassell 1976; Coates and Wortman 1980; Myers, Lindenthal, and Pepper 1975; Thoits 1983). Social support involved the existence of various ties including those of family and church, but researchers focused on the implicit and explicit benefits emanating from these ties (e.g., emotion aid, instrumental support). Third, documentation of demographic shifts in the medical care profile of health and illness in industrial and postindustrial society increased the dissatisfaction with traditional approaches. A greater chronic illness burden (as opposed to an infectious-parasitic profile) identified the career rather than the choice as the focal phenomenon, thus bringing the individual's social networks to the forefront (Clausen 1986; Haug 1981; Knipscheer and Antonucci 1990; Riley 1988). Fourth, the somewhat unexpected documentation of continued medical pluralism in both Western and non-Western nations further indicated the need for a wider conceptualization of decision making (Baer 1981; Berliner and Salmon 1980; Bhardwaj 1980; Cobb 1977).[9] Table 1 presents a fairly comprehensive, though not necessarily exhaustive, listing of the vast reserve of people all societies potentially hold who can be and are consulted during an illness episode.

This review of "alternative" themes in sociomedical research on health care raises four points similar in theoretical significance to the general critique of rational choice. First, the most obvious choice (i.e., the use or nonuse of formal medical care facilities and practitioners) does not

---

[9] Medical pluralism, characteristic of all societies, is the synchronic existence of diverse medical approaches to healing (Bhardwaj 1980; see also Leslie 1980).

TABLE 1

THE RANGE OF CHOICES FOR MEDICAL CARE AND ADVICE

Option	Advisor	Examples
Modern medical practitioners	M.D.'s, osteopaths (general practitioners, specialists), allied health professions	Physicians, psychiatrists, podiatrists, optometrists, nurses, midwives, opticians, psychologists, druggists, technicians, aides
Alternative medical practitioners..	"Traditional" healers	Faith healers, spiritualists, shamans, curanderos, diviners, herbalists, acupuncturists, bonesetters, granny midwives
	"Modern" healers	Homeopaths, chiropractors, naturopaths, nutritional consultants, holistic practitioners
Nonmedical professionals..	Social workers	
	Legal agents	Police, lawyers
	Clergymen	
	Supervisors	Bosses, teachers
Lay advisors	Family	Spouse, parents
	Neighbors	
	Friends	
	Co-workers, classmates	
Other	Self-care	Nonprescription medicines, self-examination procedures, folk remedies, health foods
None		

adequately reflect the realities of the medical marketplace or the nature of lay decision making for medical care. Second, people generally neither make a single choice nor plan a set of choices; they continue to ask advice and seek help from a wide variety of lay, professional, and semiprofessional others until the situation is resolved or options are exhausted. Third, illness triggers a dynamic, social process of coping. It is through contact with others that individuals deal with situations of medical uncertainty and find ways to solve emotional and physical problems (Kulka, Veroff, and Douvan 1979). Individuals in social networks are more than an influence on help seeking, they *are* caregivers and advisors, part of a "therapy managing group" (Janzen 1978). These habitual consultations

must be incorporated into decision-making models since studies document that the mundane options are often the most important in determining outcomes (Young 1981, p. 6). Finally, medical care systems can be conceptualized in terms of dynamic networks of interaction. Leslie (1978) makes this point forcefully: "Medical systems are generated by acts of consultation between laymen and specialists, or as concerted acts among laymen to cure, alleviate or otherwise with physical [or mental] affliction" (p. xii).

In sum, the SOS framework can draw support from the medical care arena and, in turn, has the potential to offer a more process-oriented perspective by organizing insights and providing a general conceptual base. In the SOS approach, illness careers start with an event that sets into motion a process of attempting to cope with a physical or emotional problem, given an ongoing structured system of social relations. These attempts at coping are created in negotiation with others and constrained by social structure. This orientation provides the freedom to isolate a single decision but it also emphasizes (1) ideas such as timing, spacing, duration, and order of choices (Elder 1978) and (2) that the use of official medical care practitioners, like any choice, is enmeshed in a wider pattern of help seeking. Some of the conceptual and analytic changes are possible under a rational choice framework, but the SOS framework shifts the focus to the *process* of decision making for medical care; that is, to the primacy of studying interaction and the social organization of patterned interactions directly; and, in doing so, reconceptualizes how sociodemographic contingencies influence health-care decisions by constraining or facilitating network ties (see Pescosolido [1991] for the delineation of a theoretical model of medical care choice based on an SOS framework).

A Theoretical Exposition

Adopting the social network and event-centered approach of the SOS framework leads to asking basic questions that complement those posed under rational choice models. At the same time, it exposes the current limits of adopting and pursuing a counterpart approach. There are two basic issues involved—one conceptual and one methodological. First, different orienting frameworks, as Kuhn (1970) and Merton (1957) contend, provide a base for reformulating questions and suggesting alternative ways of addressing the same issues. Conceptually, taking the SOS framework seriously, the first questions researchers adopting this approach are led to ask have not been raised in previous studies employing the dominant utilization approach. As indicated earlier, the majority of empirical studies in the dominant tradition ask: Who is most likely to use a physician, hospital, or clinic? And, under what conditions are they

likely to do so? These are the research questions that guide the rational-choice-based analysis that follows. Individuals using any other options or who do nothing are lumped together. Under the SOS perspective, the focus shifts to patterns, combinations, sets, or sequences of a wide range of lay and professional sources of aid consulted during the illness episode. The basic questions raised and pursued in the empirical exposition that follows are: Is there, in fact, a discernable set of patterns, combinations of options, or strategies that individuals use during an illness episode? And, if so, are these patterns socially organized? The sets of questions posed under the two frameworks are not contradictory; they are simply different. In the end, these differing sets of questions are complementary, providing different angles of inquiry, and ultimately, different insights on the same phenomenon.

Second, methodological issues arise. As Tuma and Hannan (1984, p. 14) note, one major obstacle to dynamic analysis is the dearth of appropriate data; here (and in general), our ability to conceptualize outstrips our ability to follow through empirically (Watkins 1980). Problems are greater for the SOS framework since most extant data, including the set used here, were organized under the dominant framework. In addition, traditional analytic strategies come from microeconomics and provide a well-traveled course for operationalizing the dependent variable and statistically examining the contingencies of decision making under the rational-choice-based framework (e.g., using OLS or logit regression for dichotomous choice). The episode approach suggested by the SOS framework, on the other hand, often calls for analyses outside standard techniques and has been difficult to take beyond empirical description (Aday et al. 1980). Under the SOS framework, even the operationalization of the dependent variable is complex and less clear-cut. The phenomenon of interest is conceptualized profoundly as a strategy, but even in the following analysis where we can only examine unique choice sets, we need to utilize an analytic technique that allows us to see the number and type of strategies that individuals employ. One way to do this is to use clustering techniques to see whether there is a limited set of unique combinations that individuals resort to over the course of a severe episode of illness (see Abbott and Hrycak [1990] for another approach to a similar problem). Even according to their developers (e.g., Hartigan 1975), these analytic tools do not have the precision of other approaches nor are they as clear-cut in their final solutions.

In addition, previous attempts to pursue alternative theoretical conceptualizations show the need to tailor methodological approaches to theoretical frameworks. Even before the incorporation of social capital into the rational choice perspective, medical sociologists faced the limits of overlaying network insights into empirical examinations based on ratio-

nal choice models. Simply appending network contacts as additional independent variables added little to overall explained variance because social characteristics share variance with network characteristics when both are conceptualized at the same level. The SOS framework suggests that, at minimum, later choices must be operationalized in conditional terms that depend on previous ones. Furthermore, the approach does not suggest the effect of mobilizing one network tie on the probability of using modern medical care practitioners but posits the central role of the structure and content of the *overall network* on the range, nature, and timing of options used.

The similarities and differences in analytic strategy called for by one or another framework as well as the potential for insights to be gained from employing both are revealed in an actual comparison. The purpose of the sections that follow is not to provide a critical test of any specific propositions of any rational-choice-based or SOS-based model but to use an available set of data (the 1975–76 National Survey of Access to Health Care) to examine the potential utility of the SOS framework premises, to illustrate the different research strategies and analytic techniques employed, and to see if taking different analytic angles on the same phenomenon leads to different insights.[10] The SOS approach requires an analytic step prior to the direct comparison of contingency effects. That is, an analysis is required to establish whether the first premise of the SOS framework, the one that remains unexplored under the rational-choice-based approach, receives support. Is there a complex but limited set of strategies that individuals use for advice and treatment during an illness episode?

The Strategies of Help Seeking

Figure 1 presents both the density and image matrices for the eight-strategy solution (i.e., seven clusters plus "none") derived under the clustering algorithm.[11] The cell entries of the density matrix represent the percentage of individuals in each pattern that choose a particular

[10] The analysis that follows is inductive because the SOS framework is an orientation to analysis and, "indispensable though these orientations are, they provide only the broadest framework for empirical inquiry" (Merton 1957, p. 88, see also Etzioni 1988, p. 17; Ordershook 1976, p. 16, Stryker 1980, p. 13). Further, the focus is on combination of options rather than time-ordered sequences.

[11] The Appendix presents details on the data, the analysis, and the interpretation of coefficients; here I include only the basics necessary for understanding the empirical results Even in this preliminary inductive exploration into the social organization of medical care contacts, data that fit the requirements to provide a useful theoretical exposition of both frameworks are rare (details on request).

	Family	Friend	Co-Worker/Classmate	Physician	Non-Prescription Drugs	Home Remedies	Druggist	None	Family	Friend	Co-Worker/Classmate	Physician	Non-Prescription Drugs	Home Remedies	Druggist	None	(N)
PATTERN 1	100	14	0	100	0	0	0	0	1	0	0	1	0	0	0	0	(506)
PATTERN 2	83	18	8	57	100	20	0	0	1	0	0	1	1	0	0	0	(65)
PATTERN 3	0	0	0	100	4	0	0	0	0	0	0	1	0	0	0	0	(302)
PATTERN 4	100	20	0	0	0	4	6	0	1	0	0	0	0	0	0	0	(66)
PATTERN 5	58	4	1	84	4	100	0	0	1	0	0	1	0	1	0	0	(71)
PATTERN 6	0	100	2	82	13	5	2	0	0	1	0	1	0	0	0	0	(77)
PATTERN 7	54	7	100	82	7	0	3	0	1	0	1	1	0	0	0	0	(68)
PATTERN 8	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	1	(42)

DENSITY MATRIX   IMAGE MATRIX

FIG. 1.—Density and image matrices for patterns of medical care contacts (N = 1,199).

medical option. A high density indicates high resort to that medical care choice in that strategy. The patterning of high densities across the row of each cluster reveals the substantive nature of the strategy. The image matrix translates these frequencies into a use or nonuse dichotomy by using an arbitrary cut-off (here, the 50% level).

These results suggest that there are, in fact, unique histories, different "cascades" of network interactions in illness episodes. As expected, because the study selected episodes on the basis of severity and recency, strategies with physicians dominate. But, to anticipate, all individuals who use physicians are not the same—their pathways are not the same, the social process that leads them to consult a physician are not the same. And, there are two patterns that do not include the use of physicians, even with this limitation. Pattern 4 (N = 66) indicates the use of the family alone for medical care advice and treatment, while pattern 8 (N = 42) represents those respondents who report doing nothing about their condition. In total, there are only three "single" strategies, where only one source of medical care was sought. In addition to those above, pattern 3 represents a single strategy, the decision-making strategy that is the focus of the rational-choice-based analyses—simply going to a physician. Respondents choosing this strategy are the individuals Freidson (1961), Ailinger (1977), and Raphael (1964) call the self-referrals. While the "physician alone" pattern represents one of the more dominant strategies (N = 302), given the severity delimiter, the data-collection method, and the existence of other strategies (which 75% of respondents use), the data in general support the utility of a conceptual framework with greater complexity of medical care choice. These results indicate that we can

gain insights by not assuming homogeneity among individuals who either seek out a physician or among those who do not.

There are, in fact, differences among respondents who seek out a physician. All remaining patterns (1, 2, and 5–7) indicate help-seeking strategies that include a physician *in combination* with some other source of advice or treatment. Pattern 1 represents the combination of consulting family and physicians and is the pattern of greatest resort ($N = 506$, 42.3%). Individuals in pattern 2 report a strategy that combines consultation of family and physicians with the use of nonprescription drugs, while those in pattern 5 seek out the same advisors but use home remedies as well. Those in pattern 7 also use family and physicians but seek out co-workers or classmates for additional medical care advice. Finally, in pattern 6, the family is not part of the therapy managing group; instead, individuals consult only friends and physicians during the illness episode.

These results do not reject the utility of a rational-choice-based approach to lay decision making for medical care nor do they support the superiority of one approach over another. Rather, they show that there are, in fact, different ways to conceptualize individuals' responses to the uncertainty of illness problems, each which provides different insights. Further, these results suggest that the SOS framework provides a real counterpart, giving additional insights into the dynamics of coping by documenting that there are different pathways to the use or nonuse of physicians.

The Social Organization of Help Seeking

The first research question requires the conceptualization and operationalization of different dependent variables for medical care under the two approaches. But the second, at this early stage of the empirical exploration using an SOS framework, shares the same fundamental structure. The basic question raised is: What are the explanatory factors at work? Or in other words, how are choices and strategies socially organized? For the rest of the analysis, we have, in essence, two dependent variables— the set of strategies derived above and the either-or decision of whether to use modern medical practitioners or not. According to the SOS framework, the fundamental mechanism at work is the operation of networks, conditioned in part by the social characteristics that tap important contextual limits on social networks. Available data do not allow a direct examination of networks; we can use social characteristics as proxies (even if not optimal) to see whether strategies are socially organized. Given this constraint, the SOS model specification reduces to the same one as in the rational choice approach with contingencies and analyses organized under the SBM (i.e., predisposing, enabling, and need charac-

teristics). This similarity, however, does allow us to see whether we can gain additional, perhaps complementary insights into decision making from the SOS approach, insights not available in analysis guided by a rational-choice-based approach.

Table 2, under "Bivariate Logit Analysis," presents the kind of analysis done routinely using a traditional, rational-choice-based framework employing a dichotomous variable (1 = used physician [85%], 2 = did not use [15%]) and logit techniques. The results are suggestive from both theoretical and policy perspectives. They indicate that blacks (twice as likely as whites), those with chronic illnesses (more than three and one-half times as likely than those with acute illnesses), and those with more severe problems (doubling their odds with each increase in severity) are significantly more likely to seek out physicians during the illness episode. Overall, the effects of social variables are quite small (and nonsignificant, even at the liberal .10 level) with size of place (b^s = 1.89), public insurance (b^s = .61), and no insurance (b^s = .79) having the largest standardized effects. In short, standard predisposing (with the exception of race) and enabling factors do not distinguish users from nonusers. Only characteristics of the illness (i.e., need) distinguish choices that suggest a very rationally based decision-making process.

In the SOS-based approach, the social organization of strategies can be assessed in the traditional regression framework. However, because the dependent variable, the strategy set, is a polytomous dependent variable (i.e., nonordered or nominal categories), multinomial logit models are the appropriate choice.[12] The second section of table 2 represents one contrast in the multinomial analysis suggested under an SOS framework and the one most like that in the rational-choice-based analysis. At first glance, the SOS approach appears to offer no additional or complementary insights into health-care decision making. Here the omitted or reference category are those individuals who report that they employed none of the many options in the survey. People with more severe conditions (as before) are more likely to use strategies involving physicians (unless

---

[12] Given the number of strategies (eight) and the number of independent variables in the final model (13), there are 28 possible coefficients (each potentially in original, unstandardized, and standardized forms with an associated t-test) with all but one providing unique information (Long and McGinnis 1981) Space limitations prohibit a complete presentation of these multivariate logit results The tables present the unstandardized effect coefficient and where applicable (i e., ordinal level variables), the standardized effect coefficient. An unstandardized coefficient, e.g., for race (black coded 1, white 0) with a value of 2 would indicate that blacks are twice as likely as whites to choose a particular strategy over the omitted strategy; a value of .5 would indicate that they are half as likely to do so The standardized coefficients allow a comparison of magnitude effects across variables.

TABLE 2

LOGIT ANALYSIS OF SOCIAL CORRELATES OF RATIONAL CHOICE (PHYSICIAN/NOT PHYSICIAN) AND SOCIAL ORGANIZATION STRATEGY (PATTERNS OF CHOICE; "NOTHING" OMITTED) CONCEPTUALIZATIONS OF HELP-SEEKING

INDEPENDENT VARIABLES	BIVARIATE LOGIT ANALYSIS. USE A PHYSICIAN (vs Not)	MULTINOMIAL LOGIT ANALYSIS (vs. NOTHING)						
		Family Only (1)	Physician Only (2)	Friend + Physician (3)	Family + Physician (4)	Family + Physician + Home Remedies (5)	Family + Physician + Non-Prescription Drugs (6)	Family + Co-worker + Physician (7)
Social location:								
Race (black)	2 20+	.84	2.15	.82	.87	1.73	.29	.99
Age	1.01 (1.16)	99 (.79)	1.01 (1.31)	.99 (.93)	.99 (.98)	.99 (.99)	1 02 (1.38)	.99 (.95)
Head's education	.92 (1.13)	1.14 (1 22)	.94 (.91)	.98 (98)	.94 (1 10)	1.05 (1 07)	1.37+ (1 62)	.83 (76)
Work status (working)...........	.93	1 42	1.77	2.51+	1.29	1.59	1.37	3.35*
Marital status (married).	94	2.09	.73	24*	1.43	.92	1.51	47+
Log of time in neighborhood	.96 (.94)	1.17 (1 29)	1.02 (1 04)	1.00 (1 00)	1.06 (1 10)	1 21 (1.37)	1 17 (1.29)	.94 (.91)
Size of place............ . ..	1.00 (1.89)	1 00 (2.15)	1.00+ (9 92)	1.00+ (11.27)	1.00 (9.67)	1.00+ (11.27)	1.00 (7 69)	1.00+ (12.02)
Skepticism of medicine	.95 (88)	87+ (.69)	.93 (.82)	.99 (99)	.87* (.69)	95 (.87)	1 05 (1 13)	.95 (.87)

1120

Enabling characteristics:

Family income	1.00 (1 06)	99 (.91)	1.00 (1 05)	99 (.66)	1 00 (1 15)	1 00 (1 03)	1 00 (1.40)	99 (.93)
Insurance:								
Public	.61	1.58	.70	53	.93	0	1.68	.57
None	79	1 67	1 01	1.17	1 30	1.80	2.27	22
Characteristics of the illness								
Condition type (chronic)	3 66*	71	2.16	2.82+	3 02*	67	1.12	1.06
Severity	1.86* (1.80)	1 09 (1.08)	2 40* (2.28)	2 27* (2 16)	2.14* (2 04)	1.56+ (1 52)	1 01 (1.01)	1 84* (1 78)
Likelihood ratio statistic	105 245**					344 26**		
df	13					84		

NOTE —All analyses are unstandardized effect coefficients Standardized effect coefficents were computed and, where appropriate are presented in parentheses All variables are coded from low to high, parentheses indicate the value coded 1 in dummy variables $N = 1, 101$
+ $P \leq 10$ (two-tailed test)
* $P \leq 05$ (two-tailed test)
** $P \leq 001$.

nonprescription drugs are included; see col. 6), but few significant effects for predisposing and enabling factors appear. To reach this conclusion, however, would be misleading. As Long and McGinnis (1981, p. 430) note, in multinomial logit analysis, individual tests of statistical significance need to be considered in combination with the overall pattern of effects (and overall significance) for a particular independent variable.

Table 3 presents the overall chi-square test for variable effects. Overall, social contingency variables are associated significantly with strategies and, most important, distinguish *among* strategies in which physicians are sought. Race, age, marital status, size of place, skepticism of medicine, and family income, as well as characteristics of the illness (both type and severity), are associated with the string of medical care options that individuals employ during a severe illness episode.

Rather than concentrating on any other specific contrast, the results are more accurately summarized through a consideration of effects across contrasts. The discussion of the effects of independent variables that follows depends heavily on the information summary provided by the graphic representations in figure 2 (based on Long 1987). Both the magni-

TABLE 3

CHI-SQUARE TESTS FOR OVERALL EFFECT OF THE SOCIAL CORRELATES OF
HELP-SEEKING STRATEGIES ACROSS ALL CONTRASTS

Independent Variables	Overall χ^2 for variable effect
Social location	
Race (black)	16.17*
Age	16.84*
Head's education	14.35
Work status (working)	10 83
Marital status (married)	52.70*
Log of time in neighborhood	6 62
Size of place	11 82+
Skepticism of medicine	21 34*
Enabling characteristics:	
Family income	12 59+
Insurance	
Public	3 85
None	9.84
Characteristics of the illness	
Condition type (chronic)	28.16*
Severity	55 36*

NOTE —All variables are coded from low to high, parentheses indicate the value coded "1" in dummy variables $N = 1,101$
+ $P \leq$ 10 (two-tailed test)
* $P \leq$ .05 (two-tailed test).

Time in Neighborhood[+]

Race*

Severity[+]

Marital Status*

No Insurance*

Public Insurance*

Condition Type*

Age[+]

Family Income[+]

Education of Household Head[+]

Skepticism of Physician
Expertise[+]

Employment*

Size of Place[+]

FIG. 2.—Graphic representations of all contrasts in multinomial logit analyses

tude distances (indicated by the place on the graph) and the significance of effects (indicated by the *absence* of lines connecting coefficient pairs) are important. The further away a coefficient effect is and the less it is connected to other effects, the more the effect of an independent variable is unique to that pattern. If coefficients are close on the graph and are connected by all possible lines, then the effect of that independent variable does not differ across contrasts.

Specifically, blacks appear to be more likely to use the physician-only strategy or a strategy that combines using physicians and friends (and to a lesser extent, a strategy including family. home remedies, and physicians). They are least likely to employ the strategy that combines the use of physicians, the family, and nonprescription drugs. Older people, on the other hand, are more likely to use this strategy that includes nonprescription drugs, but, like blacks, they are also more likely to rely only on a physician. Older people are also likely to rely on the family alone or to employ pluralist strategies that incorporate friends or co-workers. These results for older people suggest the limited nature of networks' health care advisors available to people in particular social locations.

Another group that is likely to include the use of nonprescription drugs in their help-seeking strategies are those who live in households where the head has higher levels of education. These people are also more likely to rely on home remedies in addition to family and physicians, or to rely on the family alone. They are least likely to incorporate co-workers (with family and physicians) into their medical care contacts. And while the effect of having no insurance is not significant overall, the pattern of effects is similar to that for education where the use of pluralist strategies incorporating home remedies, nonprescription drugs, or the family alone are more likely and the use of co-workers less likely. This is interesting because the latter may reflect network opportunity but the former may indicate preference.

The results continue to support the importance of social network opportunity structure in the help-seeking process. People who are working are more likely to use co-workers (in addition to the family and physicians) or friends (in addition to physicians) and are less likely to report that they do nothing about their condition. Married individuals report medical care contacts that rely heavily on the family. They are more likely to use the family alone, the family in combination with physicians, or both family and physicians combined with nonprescription drugs. They are least likely to use strategies that incorporate friends or co-workers into the therapy managing group or to rely simply on a physician.

Finally, people who report a greater skepticism of physician expertise are more likely to resort to the use of nonprescription drugs in combination with family and physicians and are more likely to use any strategy

than simply consult family members or use family and physicians. A somewhat similar pattern is found for those with higher family incomes who are more likely to employ the strategy with nonprescription drugs (also to consult friends and physicians) and less likely to rely solely on family members. Those who live in larger places are least likely to do nothing or simply consult family members, and, there appears to be a greater willingness to widen the circle of medical care advisors to include friends and co-workers (and home remedies).

In this analysis (as in the bivariate logit model), the character of the health problem is associated with utilization. Here, we are provided with much more information. Individuals with chronic or severe illnesses are more likely to employ strategies with friends (in combination with physicians), family members (in combination with physicians), and the physician directly. They are less likely to rely on home remedies or the family alone, and those with severe illnesses are least likely to do nothing or employ strategies incorporating nonprescription drugs.

These results provide a wealth of information about medical care contacts over the illness episode, information that is both more detailed, different from, but complementary to that provided by the analysis under the dominant framework. One of the most important results lies in the meaning and role of social life in help seeking. We begin to see how and where social characteristics play a role in the differential use of strategies. Social factors, more or less dismissed in the results of the rational-choice-based analysis, affect pathways to care. The effects of social characteristics in these pluralist strategies that include physicians differ in sign and magnitude in comparison with the physician-only strategy (details on request). For example, while age, income, and work status do not distinguish those who consult physicians from those who do, they do separate individuals who employ different strategies which, sooner or later, include consultation with a physician. The results do not support the sole sufficiency of a physician–no physician conceptualization, even if our goal is to understand when and why individuals seek out a physician. In sum, the findings from the multinomial and clustering analyses considered simultaneously provide support for the utility of an SOS framework for understanding medical care decision making.

CONCLUSION

At this point in the development of social science, and sociology in particular, it is no longer realistic to maintain the dichotomies that have been the source of traditional battlelines if we are to attend to the task of explaining social phenomena. Debates within sociology on the superiority of qualitative or quantitative methods, macro or micro levels of explana-

tion, static or dynamic models, to name a few, are now acknowledged to lead to unfruitful and irresolvable discussion (Bourdieu 1981; Hilbert 1990; Maines 1982; White and Eccles 1986). As Collins (1987, p. 180) notes, "we cannot live forever by merely repeating the same metatheoretical points and lambasting the same enemies." This article has addressed the movement in the social sciences away from these arid debates and toward transdisciplinary, multilevel frameworks for understanding social action. On the forefront of development lies a theoretical orientation based on an individual rational-choice-based foundation. Its wide adoption suggests, according to Swedberg (1989), a convergence toward the preeminence of a revised rational choice perspective as the basis of an explanatory framework for social science (see also Friedman and Hechter 1988; Lindenberg 1985). My purpose is not to heat up the debate about the "new Methodenstreit" (Swedberg 1989); but I do not believe that attempts to forge a synthetic framework will be represented by one, and only one, "correct mega-theory."

I have proposed a counterpart, equally synthetic and equally balanced in incorporating ideas from various social sciences to the dominant rational-choice-based orientation. This counterpart emphasizes the particular contributions and insights of sociology to form the base of an orienting framework and then profits from the insights of rational choice theory. The SOS framework begins with sociological premises (the primacy of social interaction and the nature of the situation) and overlays economic insights (utility maximization, cost-benefit analysis, purposive action). To do so does not reduce our explanation to the irrational or idiosyncratic. Rather, it shifts the focus to the sociosyncratic where rationality, purposive action, and cost-benefit calculation take place within systematically structured patterns of network interaction. The basic dynamic and starting point of the framework is social interaction (rather than individual action) with socially patterned networks of interaction producing systematic structures that link micro and macro levels. Events occur that call for action to resolve a particular situation—these career streams piggyback on the ongoing system of social life. As such, they are intertwined in social life, drawing both on affect and rationality to drive social action.

The utility of this counterpart is illustrated in the field of medical sociology where a strong interdisciplinary tradition and larger changes in social conditions have allowed greater cross-fertilization of ideas and have fueled the development of both a synthetic framework as well as alternative approaches. Taken seriously in a theoretical exposition, the SOS approach offers a different but complementary set of research questions requiring the use of more unusual statistical techniques. Even with the data limitations, the results presented here indicate that adopting the

SOS approach can provide additional insights into individuals' medical care decision making. In particular, it leads to a more refined understanding of the role of social factors in the process of coping with illness. In short, these findings support the utility of pursuing a counterpart orienting framework. Help-seeking strategies appear to be socially organized into embedded strategies. That is, (1) there appears to be a limited repertoire of patterns of choice, most of them pluralistic and some single, and (2) individuals appear to be limited in their strategies by social and medical contingencies that affect network opportunity and set the stage for the illness event.

Shifting to a new focus, as Hannan and Freeman (1986) point out in their work on organizations, opens exciting research possibilities but also poses new challenges. Stryker (1980, p. 9) strongly suggests the need to go beyond frames of reference to more precise relations among important concepts. The SOS framework presented here provides only broad, general guides. The elaboration of specific theories will require empirical descriptions of network formation and operation across diverse contexts. We have not yet developed dynamic, multimethod data collection protocols to capture both social processes and network influence (Alba 1981; Tilly 1984, pp. 26 ff.; Watkins 1980). This article represents an effort both to guide the construction of theories and to indicate the changes in data collection and empirical analysis (whether qualitative or quantitative) that should result from adopting this orientation to social action. This agenda, while complex and difficult, would move us closer to incorporating distinct contributions of sociology into multidisciplinary efforts to understand social action. But I argue in this article that it is well worth the effort to develop such theories and methods, and the theoretical exposition presented supports this contention.

APPENDIX

Detail on Data, Methods, and Measures

The National Survey of Access to Medical Care, 1975–76 (ICPSR 7730) was collected by the Center for Health Administration Studies, University of Chicago, under the auspices of the National Opinion Research Center (Aday et al. 1980, esp. chap. 6; Andersen and Aday 1980). The study is based on a household survey of the entire U.S. noninstitutionalized population (85% response rate). On the advice of one of the project's principal investigators (Aday 1985), this analysis focuses on the self-weighting main subsample of adults (i.e., $N = 1,199$ in the "illness episode" section not including data on any of the oversamples [Aday et al. 1980]). For a selected condition, individuals were asked to report

(1) in sequence, whom they talked to or contacted about their illness and what was recommended (2) in checklist form, if they talked to a set of other practitioners (e.g., curanderos, chiropractors), and (3) whether they used any home remedies (e.g., over-the-counter medicines, homemade ingested or applied medicine). Any conditions discovered by a physician in the course of a routine preventive examination (i.e., not representing a decision for an illness problem) are excluded. The data set does present limits, however (e.g., no temporal sequence; a focus on severe illnesses). Since the study is retrospective, the analysis relies on relatively stable sociodemographic characteristics and respondent information provided specifically on the illness episode (e.g., severity [Frank and Kamlet 1989]). In total, the analysis presented here provides a conservative estimate of the utility of a social-organization-strategy approach to medical care choice (details on request).

To examine the patterns of medical care decision making, all options for the focal subsample were constructed so that individuals were scored as using (coded 1) or not using (coded 0) that option (see table A1). For stability in the clustering analysis, at least 10 individuals must have selected an option. To explore the social structuring of these patterns, a number of variables were run in preliminary analyses (details on request). To increase the stability of the coefficients and to decrease the complexity

TABLE A1

USE OF MEDICAL CARE ADVISORS AND TREATMENT FOR MOST SEVERE ILLNESS
EPISODE DURING THE LAST YEAR

	FREQUENCY OF CONTACT	
SOURCE OF CARE	N	%
Physician	1,024	85.4
Druggist	10	8
Podiatrist	4	3
Chiropractor	4	3
Social worker	1	.1
Psychologist	1	1
Co-worker, classmate	76	6.3
Friend	180	15 0
Family	704	58 7
Nonprescription drugs	95	7 9
Home remedies	91	7 6
None	40	3 3
N (total)	1,199	.

SOURCE —NSAMC study, 1975–76
NOTE —Each cell in last column represents the percentage of the total number of respondents using each source of advice and/or treatment.

of interpretation in multinomial logit framework, the final model specification was determined by dropping variables that either had no effect across any pattern of medical care choice or had produced no change on the effects of other independent variables. Table A2 provides detail on the coding of independent variables in this final model.

To cluster the data, we begin with an $M \times N$ nonsymmetrical matrix of the type presented in figure A1 ($M = 1,199; N = 7$). The M represents the rows that are the number of adults who report a severe illness, in the main subsample. The N represents the columns of the eight medical care options that were selected for the clustering analysis. As figure A1 indicates, each row of the matrix represents an individual and his or her

TABLE A2

DETAIL ON CODING OF VARIABLES USED IN THE FINAL MODEL

Variable	Coding
Race .	Dummy variable 0 = white; 1 = black
Age.	18–87 years
Head education .	Eight categories from low to high attainment
Employment .	Dummy variable 0 = not working, 1 = working
Marital status ..	Dummy variable. 0 = not married; 1 = married
Time in neighborhood ...	Log of the number of months the respondent has lived in present neighborhood
Family income	In dollars from low to high incomes
Insurance status	Dummy variable set (private insurance omitted)
Public insurance	1 = public insurance; 0 = otherwise
No insurance ...	1 = uninsured, 0 = otherwise
Skepticism of physician expertise.	Likert scale of four items from "Health Opinions" section of the survey (questions and answers in nos A1, A2, A4, and A5; see Aday et al 1980, p. 402) Higher score means greater skepticism of efficacy of modern medical care.
Condition type .	Collapsed version of the first "condition" code for the health problem selected for the episode section of the survey· 0 = acute mental or physical illness, 1 = chronic mental or physical illness
Severity ..	Physician judgment of the severity of the first reported condition of the illness episode. 1 (low) to 4 (greater) severity
Size.	Census Bureau categorization of size of respondent's community from small to large

SUBJECT NO.	Family	Friend	Co-worker/Classmate	Physician	Non-Prescription Drugs	Home Remedies	Druggist
1	1	0	0	0	0	0	0
2	0	0	1	1	0	0	0
3	1	1	0	0	0	1	0
4	0	1	0	0	1	0	0
5	1	0	0	1	0	0	0
6	0	0	0	0	1	1	0
7	1	0	0	1	0	0	1
8	0	0	1	0	0	0	0
9	0	0	0	1	1	0	1
10	0	0	0	0	0	1	0
.	1	0	0	1	0	0	1
.							
.							
.							
(N)							

FIG. A1.—Hypothetical data matrix for the clustering analysis

set of choices during the episode. So, for example, the first person con-sulted only his or her family for advice and treatment.

The clustering method (FASTCLUS, SAS 1985) is designed for disjoint clustering (i.e., nonhierarchical) of large data sets and clusters on the basis of Euclidian distances. The Cubic Clustering Criterion (CCC) rep-resents an approximation of the expected value of the within-sum of squares. Values greater than three indicate structure in the data (i.e., a nonrandom set of clusters). To determine the "best" cluster set, a range of partitions (2–30 clusters) is run. Plotting the cluster values against the number of clusters run produces a graph, ideally showing a rise in the CCC, followed by a decline. However, if the data have a hierarchical structure, there may be several peaks in the graph. Such results would suggest that data are "grainy" and here, a combination of parsimony, checks for "chaining off" of small groups, as well as the CCC, is used
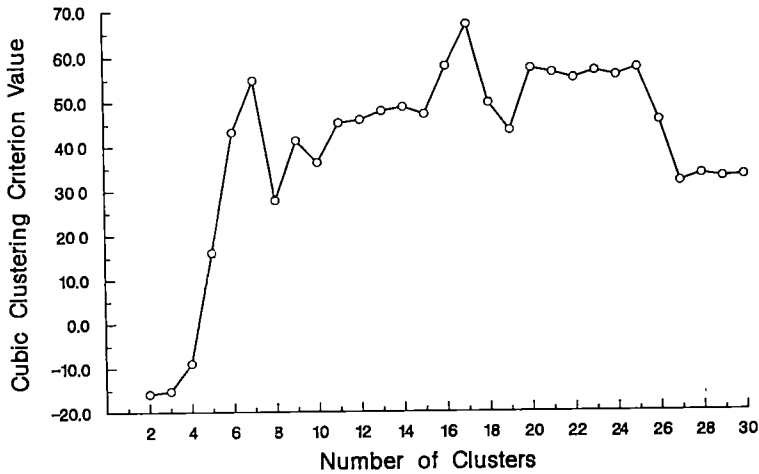
FIG. A2.—Values of the cubic clustering criterion by number of clusters specified.

to make the final decision. Here, as seen in figure A2 the seven-cluster solution (with the eighth cluster, "none," added later) represented the first peak on the graph. While it appears that the data might have a hierarchical structure, a comparison of cluster solutions with CCC values in the same range (i.e., 16, 17) indicates the basic soundness of the seven-cluster solution. Respondents who report that they did nothing about their health problem are held out as a separate vector and attached later to the pattern set for the multivariate analysis (due to its collinearity with the set of other vectors).

The density matrix is produced by computing the percentage of respondents in that pattern who select that option. It is calculated by:

$$\frac{\text{(the number of } i\text{'s in cluster C)}}{M_c J_c}. \tag{A1}$$

For the multinomial logit model, the form is:

$$P_i = e^{X\beta_i} \Big/ \sum_{j=1}^{J} e^{X\beta_j}, \tag{A2}$$

where X represents a vector of independent variables and the constraints are represented by

$$P_1 + P_2 + P_3 + \ldots P_n = 1, \tag{A3}$$

1131

where β_i is a vector of parameters that satisfy the following condition:

$$\sum_{i=1}^{J} \beta_i = 0 \qquad (A4)$$

Parameters are estimated by maximum-likelihood techniques that assess the effects of the independent variables on the log of the odds of being in one category or another of the patterns of medical care (using CATMOD in SAS 1985). Estimates are computed by comparing the effect of being in each of the patterns as opposed to being in a chosen omitted pattern (usually "none"). Only the parameter estimates that indicate the log likelihood of being in patterns 1, 2, 3 . . . $N - 1$ versus pattern N are given by the program; others can be derived (see Hanushek and Jackson 1977, pp. 210–14; McFadden 1974).

The overall fit of the model is provided by a difference of likelihood ratio statistics that is distributed as chi square with degrees of freedom equal to the number of estimated parameters and defined by 2(log likelihood at convergence − log likelihood at zero). There are a number of coefficients that can be presented. The program produces a multinomial logit coefficient of limited interpretive value and not presented here (see Long 1987). Rather, the individual effect coefficient is calculated, which is interpreted as how a unit increase in the independent variable affects the odds of respondents selecting a particular strategy versus the omitted strategy. Where applicable (where independent variables are, at least, measured at the ordinal level), a standardized coefficient can be used to compare effect magnitudes. There are two tests of statistical significance for independent variables, one indicating whether the independent variable has an effect across all comparisons and an asymptotic t-test indicating whether the variable significantly affects the choice between a particular strategy and the omitted one (Aldrich and Nelson 1984; Long 1987).

A few examples might serve to clarify the interpretation of the graphic presentation. When we look at figure 2, the graphic presentation indicates that "time in the neighborhood" does not appear to discriminate choices. All the coefficient effects are close together and the differences between them are not significant (indicated by lines between all pairs of coefficients). "Condition type" provides an example in which the independent variable does discriminate. The contrasts involving patterns 1, 6, and 3 tend to cluster together on the right-hand side of the graph and show a different effect for chronic illness from other strategies (indicated by the relative lack of lines between these and other strategies). The clustering of 4, 5, and 8 on the left-hand side of the graph (with the lines between them suggesting they hang together, i.e., are nonsignificantly different from each other) indicates choices that individuals with chronic condi-

tions are less likely to make. That is, they are less likely to employ strategies with home remedies, family alone, or doing nothing.

REFERENCES

Abbott, Andrew. 1988. "Transcending General Linear Reality." *Sociological Theory* 6:169–86

Abbott, Andrew, and Alexandra Hrycak. 1990. "Measuring Resemblance in Sequence Data· An Optimal Matching of Musicians' Careers " *American Journal of Sociology* 96.144–85.

Aday, LuAnn. 1972. *The Utilization of Health Services: Indices and Correlates.* Washington, D.C : Department of Health, Education and Welfare.

———. 1985. Personal communication with the author.

Aday, LuAnn, Ronald Andersen, and Gretchen V. Fleming. 1980. *Health Care in the U S.-Equitable for Whom?* Beverly Hills, Calif.: Sage.

Ailinger, Rita. 1977. "A Study of Illness Referral in a Spanish-speaking Community " *Nursing Research* 26:53–56.

Alba, Richard D. 1981. "Taking Stock of Network Analysis: A Decade's Results." *Research in the Sociology of Organizations* 1.39–74.

Aldrich, John H , and Forrest D Nelson 1984. *Linear Probability, Logit, and Probit Models.* Beverly Hills, Calif. Sage.

Alexander, Jeffrey C. 1990 "Commentary· Structure, Value and Action." *American Sociological Review* 55 339–45.

Andersen, Ronald. 1968. "A Behavioral Theory of Families' Use of Health Services." Center for Health Administration Studies, University of Chicago.

Andersen, Ronald, and LuAnn Aday. 1980 *National Survey of Access to Medical Care, 1975–76* (MRDF) ICPSR Data Set 7730. Ann Arbor, Mich.: Inter-University Consortium for Political and Social Research.

Aronson, Dan R. 1970 "Social Networks: Toward Structure or Process?" *Canadian Review of Sociology and Anthropology* 7.258–68.

Baer, Hans. 1981. "The Organizational Rejuvenation of Osteopathy: A Reflection of the Decline of Professional Dominance in Medicine." *Social Science and Medicine* 15A.701–12.

Berliner, Howard, and J. Warren Salmon. 1980. "The Holistic Alternative to Scientific Medicine: History and Analysis." *International Journal of Health Services* 10:133–47.

Bhardwaj, S. M. 1980. "Medical Pluralism and Homeopathy: A Geographic Perspective " *Social Science and Medicine* 14B:209–16.

Boudon, Raymond. 1988. "Will Sociology Ever Be a *Normal Science?*" *Theory and Society* 17:747–71.

Bourdieu, Pierre. 1981. "Men and Machines." Chap 11 in *Advances in Social Theory and Methodology,* edited by Karin D. Knorr-Cetina and Aaron V Cicourel. Boston: Routledge & Kegan Paul.

Burt, Ronald S 1982 *Toward a Structural Theory of Action.* New York. Academic Press.

Camic, Charles 1986 "The Matter of Habit." *American Journal of Sociology* 91:1039–87.

Cassell, John. 1976. "The Contribution of the Social Environment to Host Resistance." *American Journal of Epidemiology* 104.107–23.

Cicourel, Aaron V 1981. "Notes on the Integration of Micro- and Macro-levels of Analysis." Chap. 1 in *Advances in Social Theory and Methodology,* edited by Karin D Knorr-Cetina and Aaron V. Cicourel Boston: Routledge & Kegan Paul.

Clausen, John A. 1986. *The Life Course: A Sociological Perspective* Englewood Cliffs, N.J Prentice-Hall.

Clausen, John A., and Miriam Yarrow. 1955 "Paths to the Mental Hospital." *Journal of Social Issues* 11.25–33.

Coates, D , and Camille Wortman. 1980. "Depression Maintenance and Interpersonal Control." Chap. 7 in *Advances in Environmental Psychology,* edited by A Baum, Y. Epstein, and J. Siner. New York: Erlbaum

Cobb, Ann. 1977. "Pluralistic Legitimation of an Alternative Therapy System. The Case of Chiropractic." *Medical Anthropology* 4:1–23

Coleman, James S. 1986 "Social Theory, Social Research, and a Theory of Action " *American Journal of Sociology* 91 1309–35.

———. 1987. "Norms as Social Capital." Pp. 133–56 in *Economic Imperialism: The Economic Approach Applied Outside the Field of Economics,* edited by Gerald Radnitzky and Peter Bernholz New York. Paragon.

———. 1988. "Social Capital in the Creation of Human Capital." *American Journal of Sociology* 94 (suppl.): S95–S120.

——— 1990. *The Foundations of Social Theory.* Cambridge: Belknap.

Collins, Randall. 1981. "On the Microfoundations of Macrosociology " *American Journal of Sociology* 86 984–1014

——— 1986. "Is 1980s Sociology in the Doldrums?" *American Journal of Sociology* 91:1336–55.

———. 1987 "Looking Forward or Looking Back: Reply to Denzin " *American Journal of Sociology* 93:180–84.

Conrad, Peter. 1987. "The Experience of Illness. Recent and New Directions." *Research in the Sociology of Health Care* 6:1–31

Corsaro, William A 1990. "Interpretive Reproduction," chap. 1, manuscript. Indiana University, Bloomington

Corsaro, William A., and Thomas A. Rizzo. 1988. "*Discussione* and Friendship: Socialization Processes in the Peer Culture of Italian Nursery School Children." *American Sociological Review* 53:879–94.

Elder, Glen H., Jr. 1978. "Family and Life Course." Pp. 14–16 in *Transitions in the Life Course in Historical Perspective,* edited by T. K. Hareven New York. Academic Press.

Elster, Jon. 1975 *Leibniz and the Development of Economic Rationality.* Oslo: University of Oslo Press.

———. 1978 *Logic and Society.* New York. Wiley.

——— 1979. *Ulysses and the Sirens.* Cambridge: Cambridge University Press.

Eraker, Stephen A., John P Kirscht, and Marshall Becker. 1984. "Understanding and Improving Patient Compliance " *Annals of Internal Medicine* 100:258–68.

Etzioni, Amitai. 1988. *The Moral Dimension.* New York: Free Press.

Fabrega, Horatio. 1970. "Dynamics of Medical Practice in a Folk Community." *Milbank Memorial Fund Quarterly* 48:391–412.

———. 1978. "Ethnomedicine and Medical Science." *Medical Anthropology Quarterly* 2.11–24.

Falk, I. J., M. C. Klem, and N. Sinai. 1933. *The Costs of Medical Care.* Chicago: University of Chicago Press

Fine, Gary, and Sherryl Kleinman. 1983. "Networks and Meaning: An Interactionist Approach to Structure." *Symbolic Interaction* 6·97–110.

Fischer, Claude. 1982. *To Dwell among Friends.* Chicago: University of Chicago Press.

Fjellman, Stephen. 1976. "Natural and Unnatural Decision Making. A Critique of Decision Theory." *Ethos* 4:73–91.

Frank, Richard G., and Mark S Kamlet. 1989. "Determining Provider Choice for the Treatment of Mental Disorder The Role of Health and Mental Health Status." *Health Services Research* 24:83–103.

Freeman, Linton C., and A Kimball Romney. 1987. "Words, Deeds, and Social Structure." *Human Organization* 46.330–34.

Freidson, Eliot. 1961 *Patient's View of Medical Practice*. New York Russell Sage.
———. 1970a. *Profession of Medicine*. New York: Dodd, Mead
———. 1970b *Professional Dominance*. New York: Atherton
Friedman, Debra, and Michael Hechter. 1988. "The Contribution of Rational Choice Theory to Macrosociological Research " *Sociological Theory* 6 201–18.
Furstenberg, Anne Linda, and Linda J Davis 1984. "Lay Consultation of Older People." *Social Science and Medicine* 18:827–37.
Giddens, Anthony. 1984. *The Constitution of Society: Outline of the Theory of Structuration* Cambridge. Polity Press.
———. 1989. "A Reply to My Critics." Chap 12 in *Social Theory of Modern Societies Anthony Giddens and His Critics,* edited by David Held and John B Thompson. Cambridge: Cambridge University Press
Gold, Margaret 1977. "A Crisis of Identity· The Case of Medical Sociology." *Journal of Health and Social Behavior* 18:16–68
Granovetter, Mark. 1973 "The Strength of Weak Ties " *American Journal of Sociology* 78:1360–80.
——— 1985. "Economic Action and Social Structure. The Problem of Embeddedness " *American Journal of Sociology* 91:481–510.
Gray, John H. 1987. "The Economic Approach to Human Behavior. Its Prospects and Limitations." Pp 33–49 in *Economic Imperialism: The Economic Approach Applied Outside the Field of Economics,* edited by Gerard Radnitzky and Peter Bernholz. New York: Paragon.
Gurin, Gerald, Joseph Veroff, and Sheila Feld 1960. *Americans View Their Mental Health*. New York: Basic
Haines, Valerie. 1988. "Social Network Analysis, Structuration Theory and the Holism-Individualism Debate " *Social Networks* 10 157–82.
Hannan, Michael T , and John Freeman. 1986 "Where Do Organizational Forms Come From?" *Sociological Forum* 1:50–62.
Hanushek, Eric, and John Jackson. 1977. *Statistical Methods for Social Scientists* New York. Academic Press.
Hartigan, John. 1975 *Clustering Algorithms* New York: Wiley.
Haug, Marie R. 1981. *Elderly Patients and Their Doctors*. New York. Springer
Heise, David. 1979. *Understanding Events: Affect and the Social Construction of Action*. New York Cambridge University Press.
——— 1987 "Affect Control Theory: Concepts and Model." *Journal of Mathematical Sociology* 13 1–33.
——— 1989a. "Improvisation to Institutionalization." Paper presented at the annual meetings of the American Sociological Association.
———. 1989b. "Modeling Event Structures." *Journal of Mathematical Sociology* 14.139–69
Hilbert, Richard A 1990. "Ethnomethodology and the Micro-Macro Order " *American Sociological Review* 55:794–808.
Horwitz, Allan 1977. "Social Networks and Pathways to Psychiatric Treatment." *Social Forces* 56.86–105.
Huckfeldt, R. Robert. 1983. "Social Context, Social Networks and Urban Neighborhoods. Environmental Constraints on Friendship Choice." *American Journal of Sociology* 89·651–69.
Huckfeldt, Robert, and John Sprague. 1986. "Networks in Context The Social Flow of Political Information." Unpublished paper Indiana University, Bloomington.
Izmirilian, Harry. 1971. "Structural and Decision-Making Models. A Political Example." *American Anthropologist* 73.1062–74.
Janzen, John M. 1978. *The Quest for Therapy in Lower Zaire: Medical Pluralism in Lower Zaire*. Berkeley: University of California Press

Kadushin, Charles. 1966. "The Friends and Supporters of Psychotherapy On Social Circles in Urban Life." *American Sociological Review* 31.786–802.

———. 1983. "Mental Health and Interpersonal Environment: A Reexamination of Some Effects of Social Structure on Mental Health." *American Sociological Review* 48.188–98.

Kanter, Rosabeth Moss. 1977. *Men and Women of the Corporation* New York: Basic

Kemper, Theodore D., and Randall Collins 1990. "Dimensions of Microinteraction." *American Journal of Sociology* 96.32–68

Killworth, Peter D, and H Russell Bernard 1976. "Informant Accuracy in Social Network Data." *Human Organization,* pp. 269–86

Knipscheer, Kees C. P. M., and Toni C. Antonucci. 1990. *Social Network Research: Substantive Issues and Methodological Questions.* Rockland, Mass : Swets & Zeitlinger.

Knoke, David. 1990. *Political Networks. The Structural Perspective.* New York. Cambridge University Press.

Knorr-Cetina, Karin D 1981 "The Micro-sociological Challenge of Macro-sociology. Toward a Reconstruction of Social Theory and Methodology." Introduction to *Advances in Social Theory and Methodology,* edited by Karin D. Knorr-Cetina and Aaron V. Cicourel. Boston: Routledge & Kegan Paul.

Koos, Earl. 1954 *The Health of Regionville.* New York. Columbia University Press.

Kuhn, Thomas. 1970. *The Social Structure of Revolutions,* 2d ed. Chicago: University of Chicago Press.

Kulka, Richard, Joseph Veroff, and Elizabeth Douvan. 1979. "Social Class and the Use of Professional Help for Personal Problems: 1957–1976 " *Journal of Health and Social Behavior* 20:2–17.

Larson, Magali Sarfatti 1977. *The Rise of Professionalism* Berkeley University of California Press.

Laumann, Edward O., and David Knoke. 1987. *The Organizational State: Social Choice in National Policy Domains.* Madison: University of Wisconsin Press.

Leslie, Charles 1978 Foreword. Pp. xi–xv in *The Quest for Therapy in Lower Zaire: Medical Pluralism in Lower Zaire* by John M. Janzen. Berkeley· University of California Press.

———. 1980. "Medical Pluralism in World Perspective " *Social Science and Medicine* 14B.191–96.

Levin, Lowell, Alfred Katz, and Erik Holst. 1976. *Self-Care.* New York. Prodist.

Lindenberg, Siegwart. 1985. "An Assessment of the New Political Economy· Its Potential for the Social Sciences and Sociology in Particular " *Sociological Theory* 3:99–114

Long, J. Scott. 1987 "A Graphical Method for the Interpretation of Multinomial Logit Analysis " *Sociological Methods and Research* 15 420–46.

Long, J Scott, and Robert McGinnis. 1981 "Organizational Context and Scientific Productivity." *American Sociological Review* 46:422–42.

MacKinnon, Neil J., and David R. Heise. In press. "Affect Control Theory Delineation and Development." In *Theoretical Research Programs: Studies in Theory Growth,* edited by Joseph Berger and Morris Zelditch, Jr. Palo Alto, Calif.. Stanford University Press.

Macy, Michael. 1990. "Learning Theory and the Logic of Critical Mass " *American Sociological Review* 55:809–26.

Maines, David R 1982. "In Search of Mesostructure." *Urban Life* 11:267–79

Marsden, Peter V 1981 "Introducing Influence Processes into a System of Collective Decisions." *American Journal of Sociology* 86.1203–35.

———. 1987. "Core Discussion Networks of Americans." *American Sociological Review* 52:122–31.

Maurana, Cheryl A., Robert L. Eichorn, and Lynn E. Lonnquist. 1981. *The Use of Health Services: Indices and Correlates*. Washington D.C.. National Center for Health Services Research

McFadden, Daniel. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." Pp 105–42 in *Frontiers of Econometrics*, edited by Paul Zarembka. New York· Academic Press.

McKinlay, John. 1972. "Some Approaches and Problems in the Study of the Use of Services—an Overview." *Journal of Health and Social Behavior* 13:115–51.

———. 1973. "Social Networks, Lay Consultation and Help-seeking Behavior." *Social Forces* 51:275–92.

Mechanic, David. 1990. "The Role of Sociology in Health Affairs." *Health Affairs* (Spring), pp. 85–97.

Merton, Robert K. 1957. *Social Theory and Social Structure*. Glencoe, Ill.: Free Press

Mueller, Dennis C. 1976 "Public Choice: A Survey." *Journal of Economic Literature* 19 395–433.

Myers, Jerome K., Jacob J. Lindenthal, and Max P. Pepper. 1975. "Life Events, Social Integration and Psychiatric Symptomatology." *Journal of Health and Social Behavior* 16·421–27.

Ordershook, Peter C. 1986. *Game Theory and Political Theory*. Cambridge. Cambridge University Press.

Ostrom, Elinor. 1989. "Microconstitutional Change in Multiconstitutional Political Systems." *Rationality and Society* 1.11–49.

Parsons, Talcott. 1951 *The Social System*. Glencoe, Ill : Free Press.

Pescosolido, Bernice A 1986. "Migration, Medical Care Preference and the Lay Referral System: A Network Theory of Role Assimilation." *American Sociological Review* 51:523–90.

——— 1991. "Illness Careers and Network Ties: A Conceptual Model of Utilization and Compliance." Pp. 161–84 in *Advances in Medical Sociology*, vol 2. Edited by Gary Albrecht and Judith Levy. Greenwich, Conn.: JAI

Press, Irwin. 1969. "Urban Illness· Physicians, Curers and Dual Use on Bogota." *Journal of Health and Social Behavior* 10:209–18

Radnitzky, Gerard, and Peter Bernholz, eds. 1987. *Economic Imperialism: The Economic Approach Applied Outside the Field of Economics*. New York Paragon

Raphael, Edna. 1964 "Community Structure and Acceptance of Psychiatric Aid." *American Journal of Sociology* 69.340–58

Riley, Matilda White. 1979. *Aging from Birth to Death*. Boulder, Colo. Westview.

———. 1988 *Social Change and the Life Course*. Newbury Park, Calif.: Sage.

Rivers, W. H R. 1927. *Medicine, Magic and Religion*. London: Kegan Paul, Trench, Truber.

Romanucci-Ross, Lola. 1977 "The Hierarchy of Resort in Curative Practices. The Admiralty Islands, Melanesia." Pp 481–86 in *Culture, Disease and Healing*, edited by David Landy. New York Macmillan.

Rosenstock, I. M. 1966. "Why People Use Health Services." *Milbank Memorial Fund Quarterly* 44·94–106

Roth, Julius A. 1963. *Timetables* Cleveland: Bobbs-Merrill.

Rubel, Arthur S. 1966. *Across the Tracks*. Austin. University of Texas Press

Salloway, Jeffrey C , and P. B. Dillon. 1973. "A Comparison of Family Networks in Health Care Utilization." *Journal of Comparative Family Services* 4.131–42

SAS Institute, Inc. 1985. *SAS User's Guide: Statistics*, ver 5.0. Cary, N.C.· SAS Institute, Inc.

Saunders, Lyle. 1954 *Cultural Differences and Medical Care: The Case of Spanish-speaking People of the Southwest* New York: Russell Sage.

Schneider, Harold K. 1974 *Economic Man: The Anthropology of Economics* New York Free Press.

Sewell, William H , Jr , 1988 "Theory of Action, Dialectic, and History: Comment on Coleman." *American Journal of Sociology* 93:166–71.

Simmel, Georg. 1955. "The Web of Group Affiliation " Pp. 127–95 in *Conflict and the Web of Group Affiliations.* New York. Free Press

Simon, Herbert. 1976. *Administrative Behavior.* New York: Free Press.

Smith-Lovin, Lynn. 1987 "Affect Control Theory An Assessment " *Journal of Mathematical Sociology* 13.171–92.

Stoner, Bradley P. 1985. "Formal Modeling of Health Care Decisions Some Applications and Limitations." *Medical Anthropology Quarterly* 16:41–46

Stryker, Sheldon. 1980. *Symbolic Interactionism.* Menlo Park, Calif.: Benjamin/ Cummings.

———. 1987. "The Vitalization of Symbolic Interactionism." *Social Psychology Quarterly* 50:83–94.

Stryker, Sheldon, and Richard Serpe. 1983 "Toward a Theory of Family Influence in the Socialization of Children." Pp 471–91 in *Research in the Sociology of Education and Socialization, Vol. 4,* edited by Alan Kerckhoff. Greenwich, Conn . JAI.

Suchman, Edward 1964. "Sociomedical Variation among Ethnic Groups." *American Journal of Sociology* 70:319–31.

Swedberg, Richard. 1989. "Socioeconomics and the New Methodenstreit: On the Paradigmatic Struggle in Contemporary Economics." Paper presented at the Conference on Socio-Economics at the Harvard Business School

Thoits, Peggy. 1983. "Dimensions of Life Events that Influence Psychological Distress: An Evaluation and Synthesis of the Literature." Pp. 33–103 in *Psychosocial Stress,* edited by Howard Kaplan. New York: Academic Press.

Tilly, Charles. 1984. *Big Structures, Large Processes, Huge Comparisons.* New York· Russell Sage.

Tuma, Nancy Brandon, and Michael T. Hannan. 1984. *Social Dynamics.* Orlando, Fla.: Academic Press.

Twaddle, Andrew C., and Richard M. Hessler. 1977. *A Sociology of Health.* St Louis: Mosby.

Unschuld, Paul. 1976. "Western Medicine and Traditional Healing Systems: Competition, Cooperation or Integration " *Ethics in Science and Medicine* 3:1–20.

———. 1981. "Cross-cultural Relations. Medical Ethno-masochism." *Social Science and Medicine* 15A.1–2.

Veatch, Robert M. 1973. "Generalization of Expertise." *Hastings Center Studies* 1.29–40.

Watkins, Susan C. 1980. "On Measuring Transitions and Turning Points " *Historical Methods* 13.181–86.

Wellman, Barry, and Scot Wortley. 1990 "Different Strokes from Different Folks: Community Ties and Social Support " *American Journal of Sociology* 96.558–88

White, Harrison, Scott Boorman, and Ronald Brieger 1976 "Social Structure from Multiple Networks I. Blockmodels of Roles and Positions." *American Journal of Sociology* 88 135–60

White, Harrison C., and Robert G. Eccles 1986. "Control via Concentration? Political and Business Evidence." *Sociological Forum* 1.131–57.

Whitten, Norman E. 1970. "Network Analysis in Ecuador and Nova Scotia: Some Critical Remarks." *Canadian Review of Sociology and Anthropology* 7·269–80.

Wolfe, Alan 1989. *Whose Keeper? Social Science and Moral Obligation* Berkeley and Los Angeles University of California Press.

Young, James Clay. 1981. *Medical Choice in a Mexican Village.* New Brunswick, N.J.: Rutgers University Press

Book Reviews

From Provinces into Nations: Demographic Integration in Western Europe, 1870–1960. By Susan Cotts Watkins. Princeton, N.J.: Princeton University Press, 1991. Pp. xvii + 235. $42.50.

Richard Lachmann
State University of New York at Albany

Susan Watkins offers an elegant and convincing answer in *From Provinces into Nations* to the question of why the great differences in marriage, fertility, and illegitimacy between provinces and even villages within European nations of the early 19th century had largely been superseded with high levels of similarity within nations and continuing differences across nations by the middle of the 20th century. This book draws upon demographic data collected by the Princeton European Fertility Project for each province in 15 Western European nations from 1870 through 1960. Those data are uniform in their measures for each country throughout the period.

Watkins devotes much of the book to tracing national and provincial differences in fertility, marriage, and illegitimacy, and in comparing the relative and absolute rates of change in each region of Europe. She finds that the demographic transition had barely begun in most of Western Europe in 1870, and as a result the significant differences in marriage and childbearing practices were due largely to long-standing local and provincial customs concerning the economic resources couples were expected to have before marriage (which influenced age of marriage, the degree of sexual restraint before marriage, and therefore the rates of illegitimacy), and the extent and duration of breastfeeding (which was the principal cause of reduced fertility before the 20th century).

Watkins's strong evidence for the lateness of the demographic transition and the durability of local traditions challenges theories that use demographic factors to explain national differences in political and economic development in past centuries. While Watkins does not emphasize this point, her data set and her analysis of regional and national patterns of variation could be used to specify the kinds of temporal changes and spatial differences in political economy that can and cannot be explained by demographic patterns.

Watkins makes her greatest theoretical contribution by testing two competing sorts of explanations for the reduction of intranational demographic diversity from 1870 to 1960. Most models view marriage and fertility rates as amalgams of rational decisions made by individuals.

From this perspective, macrosocial change matters to the extent that market integration, state formation, and nation building homogenize local circumstances—most notably by reducing economic differences among individuals and providing common provisions of social welfare that serve to reduce individual uncertainty—and therefore reducing most individuals to the same narrowing range of rational choices that will yield similar decisions on when to marry and to bear children.

Watkins conducts an in-depth analysis of the data for France in order to test and challenge the rational choice approach. She finds that income differences are correlated with fertility and marriage in 1871, but not for 1961. The correlation for the earlier period is due almost entirely to the role of markets in initiating the demographic transition. By 1961 all French departments had passed through the demographic transition; therefore the effect of markets (reflected in income differences) no longer mattered at the macro level, and there is no evidence that differences in individuals' economic circumstances produced differences in marriage and fertility patterns.

If rational choices about individual economic conditions did not affect demographic decisions (at least in France in 1961) then how can differences be explained? Watkins proposes an alternative model: marriage and childbearing decisions are most affected by social pressures. Changes in political economies matter, not because they cause a convergence of individual conditions of economic calculus, but by widening the site of social control from local to provincial to national communities. Watkins's evidence for this thesis is persuasive if not conclusive. She finds that the proportion of non-French speakers in 1871 is the strongest predictor of variation in fertility and marriage across French departments in 1871 *and* in 1961. When Watkins groups modern-day departments according to what their place would be in pre-1789 French provinces, she finds that the relative variation across the old provinces is greater in 1961 than it was in 1871. The timing and extent of each province's integration into the French nation (measured most powerfully by the extent to which residents spoke French) turns out to be the best indicator of whether individuals were part of a national network and therefore responsive to national patterns of expectations about marriage and childbearing.

Watkins supports her thesis with suggestive discussions of the still relatively large variations in demographic behaviors in those countries still divided by language (such as Belgium and Norway), or by religious conflict (Ireland). She shows that the greater variations in southern Europe (Spain, Portugal, and Italy) in 1961 were due to those countries' later development, and in fact southern Europe approached the demographic homogeneity of northern Europe by 1980. *From Provinces into Nations* is an analytically powerful work, one that suggests the need for a deep reappraisal of causal models in demography and implies a rich research agenda for the future.

Autocracy, Modernization, and Revolution in Russia and Iran. By Tim McDaniel. Princeton, N.J.. Princeton University Press, 1991. Pp. ix + 239. $29.95.

Jeff Goodwin
New York University

This important and provocative comparative study picks up where the author's previous case study, *Autocracy, Capitalism, and Revolution in Russia* (Berkeley: University of California Press, 1988), left off. Tim McDaniel has now integrated the case of Iran into the general theoretical framework first elaborated in his earlier book, which suggests that "autocratic modernization" is a distinctive, albeit "contradictory," route to modernity, one that is essentially likely to culminate in social revolution.

McDaniel argues that autocratic modernization in both Russia and Iran prevented the development of a coherent state apparatus or consistent public policies, created a huge gap between the state and society, weakened elites as well as their capacity to lead other social groups, prevented the emergence of a cohesive "civil society," created subjects that tended to believe that society could be totally recast for the better, and unwittingly fostered the hegemony of precisely those oppositional groups with the most radical visions of social change. McDaniel concludes, in sum, that autocracy prevented the thorough modernization of Russia and Iran, while the modernization that did occur undermined autocracy.

This is a powerful analysis indeed, and it is generally supported by a truly impressive mass of historical evidence. The main problem with the book, however, is that the theoretical framework is not pushed nearly as far as it could go. To begin with, McDaniel provides evidence that the link between autocratic modernization and revolution is even stronger than he himself suggests. His position is that autocratic modernization does not produce revolutions so much as revolutionary situations or, at least, persistent radical challenges to autocracy; *actual* revolutions only happen when "situational crises" occur. Yet the crises in Russia in February 1917 and Iran in 1978, as McDaniel himself describes them, are for the most part intimately related to autocratic modernization and are not wholly exogenous events or fortuitous historical "accidents." Above all, the vulnerability of the old regimes to the *external* pressures, real and imagined, that precipitated these crises is clearly linked to their political and administrative incoherence as well as to their inability to fulfill their own modernizing agenda. Autocratic modernization, in other words, is even more consequential than McDaniel realizes. He fails, however, to link theoretically autocratic modernization to regime vulnerabilities within the international states system and global capitalist economy. McDaniel's notion of "modernization," indeed, seems to be drawn from the old, quasi-Parsonian school in which macrosocial change was not adequately situated within a larger transnational framework.

McDaniel might also have said much more about the potential applicability of his analysis to a broader array of cases than just Russia and Iran, even if he chose not to examine such cases in detail. Indeed, autocratic modernization is not only more consequential, but also more widespread than McDaniel seems willing to recognize. The analysis summarized above, in fact, seems generally applicable to the Batista regime in Cuba, the reign of Haile Selassie in Ethiopia (which McDaniel mentions in passing), the Somoza dynasty in Nicaragua, the Duvalier regime in Haiti, and other cases of what have been called "neopatrimonial" or "sultanistic" dictatorships (even including, perhaps, Ceauşescu's highly personalist "communist" regime in Romania). (Farideh Farhi has pointed out some of the similarities between the old regimes and revolutions in Iran and Nicaragua in her recent book, *States and Urban-Based Revolutions* [Urbana: University of Illinois Press, 1990]). The point is not that McDaniel should have mastered all these cases, which is asking too much of any mortal, but that he gains no theoretical leverage by defining the concept of autocratic modernization so restrictively that only two cases clearly exemplify it. But McDaniel, for whatever reasons, seems at pains to distinguish the Russian and Iranian regimes from all other autocracies, suggesting that they alone attempted "to create an urban industrial society as a prerequisite for present or future great-power status" (p. 7). But by his own reckoning, it seems to me, what is ultimately most crucial about these regimes is *not* that they aspired to be great powers, but rather "the arbitrary nature of [their] power and [their] inability to generate modern ideologies or organizations for political mobilization of the masses" (p. 7). McDaniel, in other words, could have made the very same substantive arguments about Russia and Iran, while employing a more encompassing definition of autocratic modernization, one with richer implications for future comparative research.

Finally, one wishes that McDaniel had teased out the theoretical sources and implications of his approach much more thoroughly. His earlier volume on Russia, in this respect, is much more illuminating, including the interesting argument that Trotsky and Tocqueville can and should be theoretically wed. It also seems to me that McDaniel unwittingly demonstrates the utility of "bringing the state back in," not in order to throw out class or cultural analysis with the bathwater, but precisely in order to do such analysis in a more satisfactory manner. Still, despite its overly concise theoretical apparatus and daunting historical detail, this book is certainly a "must" read for comparative, historical, and political sociologists of all stripes, not just specialists on Russia or Iran.

The Roots of Solidarity: A Political Sociology of Poland's Working-Class Democratization. By Roman Laba. Princeton, N.J.: Princeton University Press, 1991. Pp. xii + 247. $24.95.

Ewa Morawska
University of Pennsylvania

The book is about political developments within and by the Polish working class—its most *für sich* segment, to be exact, the shipyard workers on the Baltic coast—that assumed basic organizational form in the course of industrial strikes in 1970 and 1971, and reemerged in the summer of 1980, laying the foundations of the Solidarity union. But Laba's study is more than just a retelling of the already well-known Solidarity saga. The narrative has a clear and strongly stated polemical purpose: to refute what the author calls the "elite" or "Leninist" thesis in accounting for the roots of Solidarity. Namely, "the opinion among the political class of Poland [i.e., intellectuals] and the specialists [Western political and social scientists] that Solidarity occurred because of the leadership, the actions, and the theories of the intellectual opposition, and that consequently it is the intellectual opposition [KOR, the Committee for the Defense of Workers, formed in 1976] who are the main agents of the Solidarity revolution" (p. 3). In fact, Laba argues, Solidarity reversed the Leninist logic of social movements by demonstrating that "ordinary people such as workers were capable of coherent political activity . . . in effect transforming themselves from the objects to the subjects of politics through democratic action . . . without consciousness-raising or leadership by elites" (pp. 3, 11).

Laba's highlighting the role of the working class in the formation of Polish Solidarity is both good and timely, because—especially so in the West—despite Lech Walesa's international fame, Poland's intellectuals have been heard from much more often than the workers (I like Laba's notion that "cultural gatekeepers"—publicists and scholars on both sides—were responsible for this). As an attempt to redress this imbalance, the study is certainly worth reading. What detracts from this welcome contribution, however, and makes the author's arguments less persuasive than they could be, is "corrective overkill." In trying to prove that his case is beyond doubt, Laba overstates his claims, misrepresents (i.e., presents selectively to support his theses) studies cited as supposedly "pro-elite," and, generally, constructs both friend (the workers) and foe (the intellectuals) to fit his explanatory model. Such heavy-handedness in dealing with Solidarity's complex and multifaceted history is likely to annoy the specialists and may lead readers less knowledgeable of the historical details of this social movement to form opinions that are as lopsided as those the author set out to dispel.

Allow me to point out a few examples of this tendency to construe "clean" cases at the price of historical accuracy. Laba berates Polish intellectuals for viewing the working class as a passive mass of simpletons

with authoritarian tendencies. But, in fact, many of them, and especially KOR leaders such as Kuron, Michnik, Lipski, or Mikolajska, became in the 1970s fascinated—awed almost—by the workers' might (among the Warsaw dissident intelligentsia, I remember, it was a "politically correct" attitude, even to the point of collective self-denigration). As for authoritarianism—Laba denies any such tendency among the Polish working class—opinion surveys taken during the 1970s repeatedly confirmed such views (see D. Mason, *Public Opinion and Political Change in Poland* [Cambridge: Cambridge University Press, 1985]). A different example: as students of protest movements in repressive political environments well know, the participants' spirit alternately peaks and wanes over time. So it was in the Polish shipyards during the decade between the strikes of 1970–71 and those of 1980. It is not necessary, then, in order to demonstrate the creative role of the workers and their use of know-how from their previous experience in the nascent Solidarity union, to present them as unwaveringly combat-ready during the years in between. It is not true, as Laba claims in another "cleansing" interpretation, again one that shows a preference for a neat division between good and bad sides, that during the heated discussions, held behind the gates of the striking Gdansk shipyard in August 1980, about the organizational formula of the independent trade union, the KOR intellectuals pushed for a vertically centralized structure (Leninist, bad), while the workers had a vision of decentralized horizontal networks (democratic, good). In fact, as usually happens in such "hot" moments, the debate shifted with the logic of a rapidly changing situation and cannot be construed as a confrontation of preconceived sets of ideas.

The best part of the book seems to me the chapter "Sacred Politics," which shows the peculiar synthesis in Solidarity's ideology and artifactual symbolism of a nationalism strongly grounded in Roman Catholicism and of socialism (conventionally, these two components have been treated separately in sociological studies; an exception is A. Touraine's *Solidarity: Poland, 1980–81* [Cambridge: Cambridge University Press, 1982]).

To conclude: I was admittedly a bit annoyed while reading the book, mostly by Laba's unnecessary overstatements and simplifications as well as by his cavalier handling of theoretical issues in political sociology (for reasons wholly incomprehensible, he groups Lenin, Weber, and Hobsbawm as undemocratic centralists, the first explicit, the latter two "by implications," pp. 100–101). But still, I enjoyed my reading and find the book's corrective impulse refreshing.

Reclaiming the Sacred: Lay Religion and Popular Politics in Revolutionary France. By Suzanne Desan. Ithaca, N.Y.: Cornell University Press, 1990. Pp. xv + 262. $33.50.

David Zaret
Indiana University—Bloomington

Secularization is one of the more prominent themes in the social history of the French Revolution. Secularizing developments in the Revolution can be traced to long-term processes that slowly gathered momentum over the 18th century and to specific Jacobin policies in the second year of the Revolution, when dechristianization laws instituted abrupt, radical changes that accelerated the pace of secularization. Beginning in the fall of 1793, dechristianization not only disestablished the church but, by the spring of 1794, nearly succeeded in eliminating the public practice of Christianity Jacobin repression included seizing church goods and locking buildings, iconoclasm, imprisonment or banishment of recalcitrant clerics, and replacement of Catholic worship by revolutionary festivals and a calendar that enjoined worship on *décadis* and not on Sundays. Following the fall of Robespierre, the Thermidorean reaction provided a brief respite from enforcement and dechristianization policies. The fructidorean coup d'etat of 1797 instituted a second wave of repression that ended after Napoleon's 1799 coup d'etat.

Most prior studies of these revolutionary developments focus on two principal adversaries, Jacobin leaders and Catholic clerics, to the neglect of an important third party: the laity. Suzanne Desan's book is an important addition to recent work that corrects this omission. Her study of lay reaction to dechristianization in the Yonne calls into question several prevalent assumptions about ideological boundaries in the French Revolution. At the local level, lay loyalty to the outlawed rites and rituals of Catholicism often coexisted with support for the revolutionary regime. In this, the Yonne was not different from some other northern and central departments. The Yonne never experienced a sustained violent outburst against Jacobin rule comparable to the White Terror in other departments. This is the context for a central theme in Desan's study: the use of revolutionary rhetoric on liberty and popular sovereignty by local Catholic activists who opposed the dechristianization campaign.

Desan shows that popular Catholicism was not a static entity that remained unaffected by the revolutionary decade. Nowhere is this more evident than in the lay cults spawned by official repression. In response to Jacobin policies that padlocked church buildings and silenced clerics, villagers organized "white Masses," proceedings led by laymen, with many of the external forms of traditional Catholic worship. Laymen (they were nearly always men) also performed other priestly functions, for example, officiating at funerals, leading processions, and collecting money. Overall, the laity managed quite well without any priests. Desan's analysis of local petitions in support of Catholicism indicates that

villagers were more passionately attached to their buildings, bells, and processions than to their imprisoned or deported clerics. Desan draws other interesting inferences from her examination of these petitions. They received broad support from the communities in which they circulated—petitioning was neither confined to the lower classes nor instigated by village elites or clerics. In addition, the wording of these petitions points to a surprising syncretism: "As they infused their petitions and placards with revolutionary rhetoric and used the political institutions of the revolution to voice their claims, the religious activists continued to draw as well on traditional forms of protest. . . . This fusion of the religious and the political and of the traditional and the revolutionary increased the power of the Catholic revival" (p. 163).

Desan argues that lay cults were symptoms of long-term changes that were gradually detaching the laity from clerics as well as responses to specific dechristianization policies, especially those that interfered with Catholic rites and rituals. She traces long-term changes to the impact of Jansenism and the Catholic Reformation. Both developments led to a religiosity that was less clerical and hierarchical than earlier forms of prerevolutionary Catholicism. Jansenist priests in the Yonne cultivated a less mysterious, more edifying brand of religion, encouraging the laity to recite the Mass along with the priest. Lay cults in the Revolution extended this practice to saying the words without any priests. Within the Yonne, regional variations existed in the frequency of lay cults and other manifestations (riots, petitions) of popular support for the suppressed religion. Desan's not entirely persuasive explanation of these variations refers, among other things, to the "ambiguous legacy" of Jansenism that intensified lay religious initiatives, thereby leading to more support for popular Catholic resistance to Jacobin policies, but also accelerated secularizing trends, leading to less support for the embattled religion.

In addition to religious themes, Desan's study addresses issues that have animated historical studies of crowd behavior. Along with conducting white Masses and drawing up petitions, lay reactions to attacks on Catholicism included religious riots. Like petitioners, rioters came from a cross section of the communities in which they occurred, and in their riots Desan detects the same syncretism noted above in connection with petitions: "Rioters created spiritual and political authority by fusing traditional popular culture with new revolutionary techniques" (p. 167). This point modifies much of the recent literature on communal violence, for it calls attention to innovations occurring in the rioters' use of traditional symbols and rituals: "the forging of a language of religious protest was no simple appropriation of available symbols and rituals" (p 167). Desan also calls attention to the role of women in the religious riots. Women were in the forefront of religious riots—as was often the case in many other violent types of communal action in this era. Other claims related to gender advanced by Desan seem less certain. The motives of female rioters are said to have been more purely religious in nature

than the motives of male rioters, which "often seemed to reflect political rivalries and commercial interests as well as religious belief" (p. 198). In support of this claim Desan cites a secondary source—the primary sources expertly marshaled throughout the rest of her book are not invoked. Desan describes women as "the prime movers of religious riots" (p. 199) who "more often than men used violence to defend the integrity of a religious ceremony" (p. 200). They "seemed to view themselves as preordained protectors of both the spiritual and physical community. . . . They asserted their power to represent the community and to control its possessions as well as its culture" (p. 201). Why, then, was it the case that religious rioters "often crowned their rebellion with lay-led ceremonies" (p. 171) that, as noted above, were led by *men?*

Aside from these small flaws, this is a carefully crafted piece of social history that broadly revises widely held preconceptions about ideological currents in the French Revolution. The issues it raises go beyond those that might appeal only to specialists in communal action and religion; historical sociologists and sociologists of culture will also find this to be an important and provocative study.

Charisma and Control in Rajneeshpuram: The Role of Shared Values in the Creation of a Community. By Lewis F. Carter. Cambridge: Cambridge University Press, Rose Monograph Series, 1990. Pp. xxi + 316. $47.50.

Benjamin D. Zablocki
Rutgers University

Charisma and Control in Rajneeshpuram takes a subject that has been written about extensively and well by investigative reporters and shows the power of sociological imagination and sociological methods in adding to our understanding. It is one of the best in a long, rich tradition of monographic case studies of American communities. It deserves to be recognized as a classic of American empirical sociology.

What makes the argument in this book powerful is that it tells the story of a religious cult's rise and fall in precise historical detail while, at the same time, turning the cult this way and that under the sociological microscope so that we continually learn more about charismatic social movements and the subtle limits on toleration of religious practice in America. The author has a light touch with theory. He picks up and puts down theoretical tools as they are appropriate for his task at hand. This makes the book more difficult to read than if written from a single theoretical perspective. However, the advantage is that the reader never feels that the author has left common sense behind. A strong sense of plausibility propels us through the fact-tangled chapters.

The bare outlines of the story are as follows. An entrepreneurial Indian mystic develops a successful package of meditative and psychotherapeu-

tic spiritual practices. After a while, secular Indian society responds with social control techniques to prevent its spread. The movement is transplanted to Oregon hoping to take advantage of the unusually complete freedom of religion in America. It finds a market niche among yuppies seeking spiritual enlightenment and sexual freedom and is very successful for a while in recruitment and fund raising. However, the commune crosses "the line" and American society responds with its own equally effective social control mechanisms. The movement is shattered and the ex-members are left to put their lives back together.

Carter looks at the relationship between cult and host society from a deviance and social control perspective. He argues that the events at issue are best understood as conflicts in the tradition of the range wars of the last century. He examines the growth and adaptation of the cult over a roughly 20-year period largely from a resource mobilization perspective tempered by ethnomethodological curiosity. What ties these perspectives together is a point of view in which events are seen as resulting from actions that are rational in terms of the actors' own worldviews. The absence of social psychological concern leads to one minor drawback: a lack of attention to questions of motivation. There is little discussion of sex (either as motivator or social control mechanism). It is strange to find a book on the Rajneesh movement with no listing for sex in the index considering what a major role the combination of free love with spiritual discipline played in giving the movement its unique recruitment appeal. There is also a de-emphasis on the cognitive dissonance issues raised for the surviving loyal *sannyasins* by the plea bargaining and exile of the Bhagwan and the disintegration of the Oregon commune.

For the study of cults, this book demonstrates how powerful a tool resource mobilization theory can be for understanding religious social movements. It provides an antidote to a widespread sociological naïveté about cult finances. Carter documents the complexity of Rajneesh mechanisms for moving money across national borders and to, through, and from subsidiary organizations. A check upon financial chicanery in business is the fact that business executives rarely trust their subordinates. In contrast, charismatic organizations are able to rely on agents who are fully deployable, even in performing illegal tasks. This helps to explain both the economic success of cults during the time charisma is unchallenged and their vulnerability when charismatic authority is questioned.

For the study of social control, this book shows the importance of latent social control mechanisms in the law for limiting the range of action of deviant collectivities. In its analysis of the events leading up to and following the Wasco County elections of 1984, the book shows the importance of the invisible line between the tolerably deviant and the intolerably deviant. The book also illustrates the difficulty Third World immigrants may have in gauging the limits of toleration in America when this invisible line is crossed. Not just immigrants but many homegrown American social movements have come to grief when they blundered over this line.

For the study of the separation of church and state in America, Carter unmasks the latent fear of religious establishment that was an important motive in separating religion from state in this country. The book suggests that a wilderness beyond the frontier may be needed for absolute religious freedom. The utopian dream of a community *in* the larger society but not *of* it has been a recurring aspiration throughout American history. This book shows why such a dream is unlikely to be realized.

Hierarchy and Trust in Modern Mexico and Brazil. By Luis Roniger. New York: Praeger, 1990. Pp. xv + 236. $45.00.

Robert Gay
Connecticut College

The past two decades have witnessed a tremendous increase in the volume of research on patron-client relations in the social sciences. This research has greatly increased our knowledge of the many different forms that such relations take in the developing world. It has also dispelled the myth that clientelism is a static, unchanging phenomenon associated with precapitalist social formations. Luis Roniger argues however, that recent research has done little to further our analytical understanding of clientelism as a concept. According to Roniger, recent studies of patron-client relations have been concerned with simple topologies or with the polemics surrounding the debate between those who see clientelism as a form of social solidarity and those who see it as a mechanism of class control. Roniger seeks to overcome the limitations of past research by means of a comparative historical analysis of patron-client relations that is cognizant of structural, institutional, and symbolic commonalities at the macro level yet sensitive to unique combinations of political, social, and cultural variables at the micro level. Roniger claims that his emphasis on clientelism as an active (interactive) struggle over material and symbolic resources allows him to reconcile functionalist and structuralist perspectives and to explain the dynamics of patron-client relations over time.

In chapter 1 Roniger provides the reader with a concise conceptual definition of clientelism and a brief discussion of the economic, institutional, and cultural contexts in which clientelism is likely to flourish. He argues that conditions of economic dependency, hierarchical social stratification, and cultural orientations discouraging active participation in the process of social change lead to what he calls the "focalization" of trust. In chapters 2 and 3, Roniger describes how these conditions led to the emergence of clientelism in Latin America and the institutional context of patron-client relations in Mexico and Brazil.

The bulk of the empirical evidence in the book is to be found in chapters 4 and 5 where Roniger describes and accounts for geographical and historical variations in patron-client relations in the two countries. In the case of Mexico, Roniger compares and contrasts patron-client relations

in the Sierra Norte, in the highlands of the state of Jalisco, and in the state of Oaxaca. In the case of Brazil, he compares and contrasts clientelism in different regions of the Northeast, the Southeast, and the South. He then compares patron-client relations in these rural settings with those among unorganized sectors or urban society, among organized labor, and within the administration and political arena of the two countries. In both cases the analysis begins with the colonial era and ends with recent pressures for democratization.

In chapter 6 Roniger returns to the comparative themes introduced at the beginning of the book. Here he argues that while the core features of Mexican and Brazilian clientelism, in particular those associated with *caciquismo* and *coronelismo* (the most common manifestations of clientelism in Mexico and Brazil respectively) are strikingly similar; there are, however, important differences in "stability" and "solidity" that can be explained by wider institutional dynamics. According to Roniger, patron-client relations in Mexico are relatively unstable in that the position of the *cacique* is under constant threat by aspiring contenders from diverse social origins. This, he argues, has to do with postrevolutionary ideology, with the relative political autonomy and strength of collective actors and with high levels of intraelite competition. By way of contrast, patron-client relations in Brazil are relatively stable. They are less threatened from above and below and, perhaps more important, are more solidly anchored in cultural norms governing social interaction.

The idea of a "shared sociomoral model of interaction" (p. 174) as an essential element in clientelist exchange is a theme the author develops in the next chapter. In chapter 7 Roniger asks that we put aside the more simplistic evolutionary approaches to the study of clientelism that claim a distinction between informal, instrumental, short-term, or "modern" patron-client relations and those of a more "traditional" form. Instead we should recognize that "in most patron-client relations . . . specific exchanges are embedded in long-term, generalized exchanges that offer delayed or reflected benefits, generally in the political and social spheres," and that, as such, "considerations of personal power, solidarity, and the symbolic meaning of both have been relatively more important" (p. 173).

The idea that symbolic aspects of patron-client relations shape commonsense notions of reality independent of structural and institutional change is both useful and perceptive and one that has framed much of the recent debate over the process of democratization and social change in contemporary Mexico and Brazil. The idea that trust and mutual obligation are as important, if not more important, to patron-client relations than coercion and domination is less clear-cut, however, and one that will surely add fire to the polemic the author tries to avoid.

Animal Rights, Human Rights: Ecology, Economy and Ideology in the Canadian Arctic. By George Wenzel. Toronto: University of Toronto Press, 1991. Pp. ix + 206. $55.00 (cloth); $18.95 (paper).

Dona L. Davis
University of South Dakota

George Wenzel's excellent book about human and cultural survival in face of economic, political, and ideological imperialism places the Canadian seal controversy in the wider context of the history of intercultural relations between the Inuit (Eskimos) and the Qallunaat (whites) of Canada. Basing his analysis on data collected from over 22 years of ethnographic research, interviews with animal rights advocates, participation in sealing (seal hunting) and wildlife conferences, and a wide range of secondary sources, Wenzel presents a detailed examination of the ecology, economics, and ideology of sealing among the Inuit, past and present. The history and ideology of the antisealing campaign and the animal rights movement are discussed with reference to the impact of their more current political action campaigns on the Inuit. The book, however, should not be dismissed as a regional study. *Animal Rights, Human Rights* addresses some extremely topical and controversial questions such as What constitutes a people's culture and cultural identity? Do insiders or outsiders have the right to determine that identity? What are the political, social, and economic ramifications of that process in the modern nation state?

The thrust of Wenzel's argument concerns the uncompromising position taken by animal rights advocates that, because the Inuit use modern technology (snowmobiles and rifles) to harvest seal, and because they sell pelts for cash that is used to buy imported goods, the Inuit have no claim to being a traditional society. In rejecting the proposition that Inuit perception and behavior toward animals can be rooted in anything more complex than commercial exploitation, animal rights advocates have not only acted to redefine Inuit culture but have also acted to determine what measures are applicable to cultural analysis. Once focused on the morality of sealing, the animal rights movement has turned to a very critical examination of Inuit culture itself. Wenzel counters that this focus on artifacts is ethnocentric, it misconstrues the nature of culture, and it fails to acknowledge native cultural reality. To make his point, Wenzel relates modern and traditional sealing practices to Inuit ecology, socioeconomics, and ideology and argues that modern artifacts have been adapted to traditional culture and not vice versa, that values related to hunting pervade all areas of Inuit life; and that, despite contact with administrators, missionaries, and developers, the Inuit are neither assimilated nor in cultural confusion.

Wenzel provides a convincing argument that modern techniques are combined with traditional knowledge, systems of sharing, and technology to make contemporary seal hunting more efficient and ecologically sound

than earlier ways of hunting. From an ecological perspective, Inuit seals are an important, abundant, accessible, and cheap component in the diet. Inuit sealing is also portrayed as central to nonmaterial components of life. Socially, sealing is tied to complex, daily dynamics of the kinship system that regulates the distribution, preparation, consumption, and celebration of the hunt and its products. The ideological aspects of hunting include the Inuit view of animals as helpmates, with many rights including the rights to refuse hunters, to be treated with respect, and to be hunted and used wisely. Additionally, the view that meat and pelts of animals are a means by which humans contribute to the collective security of the social group is complexly interwoven with Inuit notions of collective identity and community.

Inuit culture has survived for over 4,000 years, but never before, argues Wenzel, has its physical and biological base been so directly challenged. The negative consequences of the European Community's bans on the sale of pelts include loss of local autonomy, economic marginalization, overexploitation of less abundant food resources, and the cultural alienation of young men. Yet the more insidious consequences of the sealing ban lies in the fact that, reacting to the challenge posed by native rights advocates, animal rights activists have recently expanded their seal/fur protest rhetoric into much wider concerns of native peoples and are attacking nearly all aspects of Inuit wildlife use and land rights claims.

Wenzel is at his best discussing Inuit history, social economy, and ecology. His analysis focuses on the more extreme and dogmatic elements of the animal rights movement. For those not familiar with the range of different groups and their different constituents, agendas, and leaders, Wenzel's analysis can be confusing. Written with a minimum of anthropological jargon, and dealing with an issue many people feel strongly about, the book is suitable for a general readership and for undergraduate use. A good example of advocacy anthropology, the volume is also suitable for courses on ethnic minorities, culture change, cultural ecology, Fourth and First World issues, and social movements.

Women in the Sanctuary Movement. By Robin Lorentzen. Philadelphia: Temple University Press, 1991. Pp. 229. $34.95.

Steven M. Buechler
Mankato State University

This ethnography is based on in-depth interviews with 29 laywomen, clergywomen, and nuns who participated in the Chicago sanctuary movement between 1982 and 1987. These interview data are used to paint a detailed picture of the sanctuary movement and explore larger questions about the role of women in social movements.

Lorentzen locates the sanctuary movement within a tradition of moral

reform movements in which women have played a major role and in which religious institutions have often provided the free space and cognitive liberation needed for mobilization. After sketching the history of the Chicago movement as a diverse, multidenominational network, the book focuses on ideological splits within the movement. The major divide is between a humanitarian approach epitomized by resettlement efforts originating in Tucson and a political approach exemplified by movement-building tactics in Chicago. This regional difference overlapped with a gender difference; males in both locations favored the humanitarian approach, while some women were more likely to take a political approach. While this appears to contradict gender stereotypes, Lorentzen argues that women's prescribed role of affiliation with others led many to identify with refugees and to advocate shared power rather than humanitarian paternalism.

Despite this gender difference, the greatest ideological divide was between women. Many laywomen also favored a humanitarian approach that encouraged refugees to assimilate, while religious women developed a political approach that used the situation of the refugees to challenge U.S. foreign policy. For religious women, sanctuary work was an intentionally organized, public, political challenge to dominant power; whereas for laywomen sanctuary work was an informal and situational response to humanitarian needs.

In many sanctuary groups, women developed a feminist orientation expressed by a distinct leadership style that rejected hierarchy in favor of mutual empowerment. However, one of the more intriguing findings of this research is that there was ongoing tension between the worldviews of sanctuary workers and women refugees. Central American women were not nearly as willing to assimilate as humanitarian reformers assumed, nor were they very receptive to the feminist view of gender relations advocated by more political reformers. In this case, feminist ideologies could not overcome class and cultural differences.

In later chapters, Lorentzen describes the paths by which women became involved in the movement and the kinds of conversion experiences that mobilized them. She also analyzes differences in "life structure" to explain variations in women's participation. For many laywomen with family connections, sanctuary work reflected humanitarian concerns and limited commitments to feminism or liberation theology in the context of cross-cutting family obligations. For religious women free of such ties, activism was more deeply rooted in a feminist theology that linked various oppressed groups to the need for liberation. Because of their greater cognitive liberation from patriarchal worldviews, religious women assumed the activist role with less contradiction than did women who were also wives and mothers.

While this study is rich in anecdotal evidence, there are some larger conceptual, interpretive, and theoretical issues that are not always satisfactorily resolved. The book is initially described as a study in how women transformed sanctuary into a feminist movement (p. 5); in a later

chapter we are told that "sanctuary is not a feminist movement" (p. 175). A clearer definition of terms could clarify these contradictory claims. Overgeneralization is also a problem; many readers will question whether 29 interviews justify the claim that "this study shows that women are inclined toward democratic movements" (p. 190).

Some of the most intriguing evidence, on the other hand, is not fully utilized. Many activists espoused an essentialist view of a distinct women's nature. However, the experiences of refugee women provided an immanent critique of this view because class, culture, and nationality typically overrode gender differences for them. Yet we are told little about how sanctuary women responded to this implicit challenge to their perspective.

Finally, there is a theoretical contradiction in this analysis. To understand the role of women, Lorentzen relies on feminist theories about women's distinctive ways of knowing, being, and acting in the world. To analyze their participation in this social movement, she uses resource mobilization theory, including models of rational actors who engage in cost-benefit analysis when deciding to join social movements. The latter is precisely the kind of gendered, patriarchal theory that is typically rejected by many strands of contemporary feminist theory.

As a study of the sanctuary movement, this book makes an important contribution. As an analysis of women in social movements, it raises more questions than it answers. In both cases, it suggests promising leads for future research on these topics.

Between Feminism and Labor: The Significance of the Comparable Worth Movement. By Linda M. Blum. Berkeley: University of California Press, 1991. Pp. 249. $30.00 (cloth); $11.95 (paper).

Joan Acker
University of Oregon

The comparable worth movement, campaigns to raise women's low wages in female-dominated jobs, is now almost 20 years old. As practical experience accumulates, assessments of effects on wages, workers, and employers replace speculation about possible consequences. Linda Blum makes a valuable contribution to this growing empirical literature with accounts of the comparable worth movements in the city of San Jose and in Contra Costa County, both in California. She uses these cases to show the complex connections of gender and class interests in pay equity politics, to examine shifts in gender consciousness as the feminist workplace agenda moved from affirmative action to pay equity, and to assess both the limits and radical possibilities of comparable worth.

The 1981 San Jose comparable worth victory was the first pay equity success, a beacon for activists in Contra Costa County and many of the rest of us who were pursuing pay equity elsewhere. San Jose in the 1970s

was the "feminist capital of the world" (p. 56). City government was dominated by feminists committed to affirmative action, but, as many ordinary workers realized, this strategy could result in better jobs for only a few. Clerical women workers began to organize and to campaign for higher wages for all clerical workers rather than promotions for one or two. Joining the AFSCME local, they took it over by 1978 and placed pay equity at the top of the bargaining priority list. After a comparable worth study in which union members participated, trouble erupted as city government, including elite feminists, resisted full implementation of wage increases. A dramatic strike finally moved the city to settlement. Thus, in San Jose, a strong feminist presence in city politics at the top supported the organization of rank-and-file women, but this feminist coalition across class lines weakened as women in city leadership disagreed with low-paid union women on implementation of the comparable worth study.

While in San Jose comparable worth brought women into the union, in Contra Costa County women with a high degree of class-consciousness moved from a view of their problems as class issues to seeing them also as gender issues. Comparable worth activists in Contra Costa had a longer, more difficult, and less successful movement for pay equity than did those in San Jose because they faced an unsympathetic, male-dominated city government, a reluctant, male-dominated union, and a county economy in fiscal difficulty. Nevertheless, they made small gains during the 1980s and expected implementation of a comparable worth plan in 1991. In San Jose and Contra Costa, as in many other places, pay equity efforts "put real money in real people's pockets" (p. 130), but the amounts were modest and the wage gap remains.

What is the significance of the comparable worth movement as seen through these two cases, as well as others such as the cases of Minnesota and Oregon? Comparable worth is not just a technocratic reform crafted by elite women in a top-down process, nor does it benefit middle-class professionals the most. It can be a grass-roots movement of low-paid, pink-collar women acting in their own behalf, and such workers have been the main beneficiaries. Comparable worth is a powerful mobilizing idea for low-wage women workers because it affirms the value of their work and suggests a solution that increases respect as well as wages, addressing the needs of working women more directly than affirmative action, which denigrates the work many women do by arguing that they must enter men's jobs to improve themselves. By affirming the value of the devalued female job, comparable worth also questions the market as a neutral process of value determination. Thus, it may be the only effective challenge today to the dominant ideology of the market.

Comparable worth can also be an avenue for bringing feminism and organized labor into a new coalition. In these two cases, women's consciousness of both gender and class developed through the process of self-organization, leading to labor militancy. Although alliances with, and initiatives of, elite feminists were important, Blum warns that the

radical potential of comparable worth is only possible if it remains a grass-roots movement, because elite women may be willing to settle for a quick technocratic solution, cutting off the possibilities of new alliances, new mobilization, and transformation of political discourse.

Labor-union women in many other industrial countries are now making comparable worth demands. This signals, I believe, a fundamental shift in the organization of gender and class. Now most women must work and their pay becomes a class issue, while, at the same time, the old masculinity based on the male wage is undermined. Comparable worth can be seen as part of a new, long-term struggle around gender and class issues. This book contributes to our theoretical and empirical understanding of these processes.

Democracy without Women: Feminism and the Rise of Liberal Individualism in France. By Christine Fauré Translated by Claudia Gorbman and John Berks. Bloomington: Indiana University Press, 1991. Pp. viii + 196. $29.95.

Joan W. Scott
Institute for Advanced Study

In *Democracy without Women* Christine Fauré, a French sociologist who came of age as a feminist in 1968, argues that feminism in France owes far more to liberal individualism than to socialism. She writes against a historiographic tradition that has long associated the first use of the term "feminism" with the utopian socialist Charles Fourier and that has linked women's demands for equality to the left in various revolutionary movements beginning with the French Revolution. Instead she suggests that "the progress in society of notions of equality and the individual" are the ground for feminist struggles and that these "philosophical, religious and political 'revolutions' " (p. 15) anticipated, by many centuries, the French Revolution.

Fauré points out that feminism found articulation whenever a notion of individuality was expressed. Christine de Pizan drew on Augustinian ideas of individual introspection to affirm women's capabilities in the face of misogynist clerical attacks in the 14th century. In the 16th century, in very different terms and contexts, feminists took up humanist emphases on educational and religious reform in order to argue that women, too, were individuals whose capabilities must be developed. The Protestant Reformation provided women another source of individualist egalitarian argument, as did the doctrines of those who sought to limit the consolidating force of royal power, even though many of the Protestant and political theorists (like Calvin or Jean Bodin) argued for the exclusion of women from public and political activities. Feminists pointed out the logical contradictions in these theories that made universalist claims for the rights of genderless abstract individuals, on the one hand, and exclusions

based on gender, on the other. In a sense, the long history of feminism is the history of the attempt to resolve this contradiction by proving that women, too, are individuals with the same rights as men.

Fauré documents this history loosely and descriptively. She presents long summaries of the writings of various theorists, losing her argument in a welter of detail. Chapters fade off rather than end conclusively, comparisons that are promised (such as the one between Montesquieu, Rousseau, and Condorcet) are never fully drawn. The book is more an impressionistic intellectual history than it is a probing analysis of feminism's relationship to liberal individualism. The discussions of theorists like Bodin, Rousseau, or Condorcet add little to what is already known of their work. There are no new insights for political theorists about the canonical thinkers (as there are in Carole Pateman's provocative and original *Sexual Contract* [Stanford, Calif.: Stanford University Press, 1988]). Fauré's most significant contribution is to introduce some previously obscure feminist thinkers—articulate, clever, and witty—to the historical record.

Most disappointing is the lack of any sustained analysis of the problem presented. How was exclusion built into the very terms of liberal individualism? The kind of close, deconstructive reading of texts that this question requires is absent from the book. Instead Fauré implies that misogyny or ill-will or lack of clear thinking resulted in women's exclusion by men. The recurrence of exclusion is made to seem somehow external to liberal ideology itself, rather than intrinsic to its logic.

Feminists' demands for equality are also made to seem more straightforward and simple than they were. The problem of sexual difference— which all these women theorists had to address, and which they did address very differently—is nowhere taken up by Fauré. It was not possible simply to demand equality in the face of doctrines that justified exclusion in terms of natural or God-given differences between the sexes. Feminists had to contest those justifications. How they did it, the difficulties and paradoxes such contests created in feminist thinking, and the implications of those difficulties for feminist politics—these are the issues that need analysis by historians of feminism. They have begun to be addressed for France in the work of Geneviève Fraisse (*Muse ·de la raison: La Democratie exclusive et la différence des sexes* [Paris: Alinea, 1989]), which, although limited to the early 19th century, is far more satisfying as a sustained and serious analysis of these issues.

Democracy without Women should be read as a feminist attempt to rethink political theory outside its usual terms; but it is not ⎣he best example that can be offered of such a rethinking. It is too constrained by its contemporary French context to be accessible to most American readers and too narrowly focused to serve as a good example of what French feminist scholarship has produced. Yet it identifies a topic worth attending to and, in so doing, opens a field of inquiry that others will surely pursue.

Managing Gender: The State, the New Middle Class, and Women Workers, 1830–1930. By Desley Deacon. New York: Oxford University Press, 1989. Pp. x + 308.

Hester Eisenstein
State University of New York at Buffalo

In *Managing Gender*, Desley Deacon has written an intensely researched and complex tale: a case study of the interactions of state, class, and gender in the public service bureaucracy of New South Wales, one of the two most populous and influential states in the Australian federation (the other is Victoria). Deacon argues that in discussions of the role of the state as an agent of ruling-class interests and of patriarchy, too little attention has been paid to those who actually staff the state apparatus and carry on its day-to-day operations. It is her contention that, in the Australian context of nation building, from the 1860s to the 1900s, the "new class" of intellectuals, professionals, journalists, and lawyers created a public service in their own image. In established versions of Australia's rise to nationhood, the battle for control between capital and labor has held center stage. Deacon proposes to add the story of the new middle class, with their own interests and their own agenda. It is a story with fateful consequences for women.

Following the departure of the British colonial government, and the establishment of self-government in 1856, administration remained embryonic. The economy was largely based on family labor, and white women easily obtained important positions as teachers and postmistresses and shared the work of running farms and managing small businesses like taverns with their husbands. In a decentralized situation with local autonomy, gender was not a central employment issue.

In the 1870s and 1880s, the New South Wales public service expanded, organized in the interest of state politicians who depended on large-scale public works to win reelection. Big government departments developed, staffed by such male technical personnel as engineers. But the uneven development of the service ensured that areas of employment like teaching and the postal service remained open to women. Political patronage flourished, and women benefited as well as men.

Late in the century, a new phase opened as the "new class" of public servants presided over the professionalization and reorganization of the service as a whole. Deacon characterizes the new class, following Alvin Gouldner and others, as "a third class standing rather uneasily between the bourgeoisie and the proletariat. . . . This . . . new middle class can be thought of in general terms as consisting of those workers who depend on the sale of educational, technical and social skills, or 'cultural capital' " (p. 4). They are exemplified in Deacon's account by the powerful figure of T. A. Coghlan, who brought the new science of statistics to bear in pressing for his conceptions of how the state should regulate the labor

market and what the role of wives and mothers should be in the "state socialism" of 20th-century Australia.

An "ambivalent champion" of women, Coghlan allied himself with the generation of feminists—themselves members of the "new class"—who succeeded in winning female suffrage in 1902 and who, for a short period of time, gained equal opportunity, at least for unmarried women, in the public service. Deacon lets us hear the surprisingly modern voices of these women, who are self-consciously aware of their pioneering status.

But Coghlan and his allies were pushed aside by a new coalition of forces, as the modern configuration of Australian politics emerged, divided between (social democratic) Labor and the (conservative) ancestor of the contemporary Liberal-National coalition. Cost cutting became the impetus for instituting discriminatory practices against women. The dual labor market was enshrined in law by the new industrial arbitration system, which instituted a family wage, and fixed women's wages at the level of the notorious "two-thirds of a man." The salaries for public service workers followed this model, limiting women's income and annual increases, virtually ensuring a high turnover of women employees. Along with government-sponsored campaigns for professionalized infant care and for an increase in the white birthrate, which stressed motherhood over waged work for women, the unequal wage system succeeded in pushing women out of permanent careers in the public service and in discouraging them from entry into public life more generally for several generations. (Deacon's story ends in 1930 and thus does not cover the period in which this trend is reversed with the lifting of the "marriage bar" in the public service in 1966, and the introduction of equal pay onto the books, if not in practice, in 1972.)

The value of this study is in its careful documentation of several turning points in the establishment of a particular gender order in one white settler society, illustrating the fact that these events are the product of historical actors and not of some larger inevitable machinery. As Deacon notes, "The ideas and practices of new middle class men from the 1890s to the 1930s concerning women as workers, and their power to transform them into accepted conventions of thought and behaviour, were the product of a particular time and place and cannot be attributed to some essential quality of either men or the state" (p. 219).

The book is dense and detailed and assumes some knowledge of Australian political history. It also makes some leaps. Coghlan's views are presented as virtually representing those of his class. Even if we accept this, does it mean that the professionalization and consolidation of a professional public service based on ideas of merit was not in the long-run interest of the capitalist class? But the questions raised illustrate the virtues of this excellent study, which provides rich food for thought about the historical creation of an enduring class, race, and gender order that many of us are currently struggling to dismantle.

Dependence and Autonomy. Women's Employment and the Family in Calcutta. By Hilary Standing. New York: Routledge, Chapman & Hall 1991. Pp. x + 198. $16.95 (paper).

Virginia S. Fink
University of Virginia

Hilary Standing describes the tightly articulated connections between economic work, gendered roles, and the family in her ethnographic account of women in Calcutta. In the process of reading you become aware of the pervasive social creation and recreation of women's dependence and their occasional ability to achieve a degree of autonomy. Focusing her analysis on changing household income patterns, she illustrates the complex interaction of the gender and employment roles of women in a variety of traditional Indian families. Her use of historic statistics from the 1911 Indian census record is effective in showing the changes and stability in the gendered work roles for this large city.

Deconstructing the broad sociological categories of family and households allows a better understanding of women's lives. We also see how autonomy and dependence change over a woman's lifetime. Often the very concepts of household income and head of household have obscured the dynamics of women's lives. The understanding of the social creation of female dependence is lost in attempts to understand women's lives by simply quantifying women's work. In changing the focus of her analysis, women emerge as human actors living in rapidly changing times. They are structuring their world within a system of shared meaning about work, gendered roles, and family. For many it is a world where some types of dependence seem very normal, where governmental decisions about forms of transportation and the availability of child care can prevent women's autonomy more than their small salaries.

Some of the insight gained from reading comes from the comparison of a diversity of experiences. The range of situations enables the reader to grasp the broader normative ideological patterns the women all operate in. Using a representative sample of 114 households in Calcutta, India, Standing interviewed Hindu, Islam, and Christian women from Brahman, Kayastha, Sudra, and Baidya castes.

The book aptly illustrates the importance of understanding how the cultural norms of the society structure female behavior and change. The culture offers a basic "tool box" of strategies. To use Ann Swidler's terms (see "Culture in Action: Symbols and Strategies," *American Sociological Review* 51 [1986]), these tools often limit the set of skills that women bring to the marketplace and also limit the working behavior of women in Calcutta. This is an interesting analysis of women's dependence in developing countries because the author stresses the difference between an *autonomy of individualism* that is an aspect of autonomy in the Western world and autonomy within a familial based society such as India.

1160

It also highlights many of the same decision-making processes made by Western women who have also strived to balance family, work, and identity. This book holds up a mirror so that we can see these processes more clearly in our lives.

Occasionally I would have liked more detailed labeling for some of her tables, but this flaw hardly offsets the richness of the text and the discussion of the intertwining of women's roles in the domestic arena and their work as wage laborers.

While this seems an excellent book for anyone contemplating fieldwork in India or even other Third World countries, it will also be a good text for a course on development, women and work, or the family.

Standing ends with policy suggestions that seem to reflect the needs articulated by the women she describes, and, while the suggestions echo the voices of women in many cultures, they would all seem to increase the autonomy of women in Calcutta.

A Future for the American Economy: The Social Market. By Severyn T. Bruyn. Stanford, Calif.: Stanford University Press, 1991. Pp. 424. $45.00.

Richard Swedberg
University of Stockholm, Sweden

At one point in *A Future for the American Economy,* the following story is told. Frank Knight was once asked what constituted "the most important economic problem" in the United States. He answered, "the question has no definite answer. Most problems involve some use of means—hence [they] demand 'economizing,' avoiding waste and futility. Accordingly, economic problems form no distinct class, and any list would be largely arbitrary" (p. 19). Within the context of Bruyn's book, this story is told to illustrate the failure of the economists to come to terms with the many ills that are currently besetting the American economy. We are told that a totally new way of thinking about the economy is needed, if things are to improve. This book is Bruyn's own attempt to provide us with such a new way of thinking.

Maybe "vision" is the most appropriate word for what we find here. According to the author, the U.S. economy is already, by and of itself, moving in the direction of a self-regulating and equitable kind of economic system in which the state will play a minimal role. The end result of this process is an economy that is governed through an intricate system of mutually balancing and internally self-governed federations of trade associations and "associational" trade unions. It is imperative, Bruyn says, that social scientists help to clear the way for this new type of economy through socioeconomic studies.

A Future for the American Economy is divided into three major parts. There is also a concluding section as well as an extra chapter (wisely

relegated to an appendix) that extends the argument to the international economy. In part 1, Bruyn presents the theoretical part of his argument, which he describes as having "an analytical dimension and a normative dimension" (p. 6). The emphasis is on evolution in a Hegelian-dialectical manner: what is emerging around us is neither capitalism nor socialism, but a synthesis of the two—what the author calls "the social market." Most of us only notice the conventional side of the economy (the manifest economy), but there is also another, emerging one (the latent economy); it is the latter that the book wants to draw attention to.

In part 2, Bruyn looks at this latent economy. He discusses workers' self-management, social investment, nonprofit organizations, and the service economy. All of these are said to be indications that the U.S. economy is slowly but surely moving toward a self-regulating type of economic system in which equity and efficiency will be present. The author in particular pays homage to self-managed firms. These are said to be very productive and efficient; they have less absenteeism and less pilferage than ordinary firms, and so on.

In part 3, the author discusses the role that the corporations are to play in "the social market." They are basically to be organized in federations of self-governing trade associations. Again, since reality is moving in this direction, today's trade associations already display some of their liberating potential. Still, much remains to be done on this score, as Bruyn acknowledges. The author also emphasizes that social scientists must do what they can to support this development. Bruyn, for example, envisions a kind of Council of Social Advisors, to match the Council on Economic Advisors. He also suggests that we need to construct various "social indexes," which will tell us whether the trade associations have become truly self-regulating and democratic.

In my opinion, there are many positive qualities to this book. It contains, for example, fine sections on social investment and self-managed firms. It is also refreshing, for a change, to read an analysis that emphasizes the role of cooperation rather than competition in the economy. In the final analysis, however, it is difficult to know what to do with the mixture of normative and analytical elements that is so characteristic of this work. Maybe it is fitting at this point in the review to refer to the well-known story that Albert O. Hirschman tells in a similar context in *The Passions and the Interests* (Princeton University Press, 1977), namely apropos Montesquieu's inspired but faulty theory that industrial capitalism would make society more pleasant and peaceful (*"doux commerce"*). The story goes like this. The rabbi in Krakow had a vision that the rabbi in Warsaw had suddenly died, and his congregation was very impressed by this extraordinary display of visionary power. Some time later, however, a few people from Krakow traveled to Warsaw and found that the local rabbi was very much alive. On their return, they told what they had seen. A few staunch supporters of the rabbi in Krakow were, however, not impressed. The details in the original vision, they conceded, might have been wrong—"Nevertheless, what a vision!"

Antitrust and the Triumph of Economics: Institutions, Expertise and Policy Change. By Marc Allen Eisner. Chapel Hill: University of North Carolina Press, 1991. Pp. xii + 302. $34.95.

Robin Stryker
University of Iowa

Antitrust and the Triumph of Economics, by Marc Allen Eisner, solidly contributes to the political-institutional approach to policy-making. Having examined the enforcement processes and records of the Federal Trade Commission and the Antitrust Division of the Department of Justice from these agencies' creation in the early 20th century to the present day, Eisner argues that the presumed Reagan revolution, which used Chicago school economics to justify negating most traditional goals of antitrust, was no revolution at all. It was an evolutionary and, to some extent, unintended product of prior organizational changes fusing law and economics into an intellectual and institutional framework dominated by economic expertise.

When economics became ascendant at the antitrust agencies, the paradigm of economic structuralism dominated industrial organization economics and the antitrust "community of expertise" (p. 92). The latter is both a network of actors who critically examine policies and chart alternatives and "a common body of knowledge [and] analytic tools" (p. 92). Economics came to dominate antitrust enforcement for various reasons. Antitrust had low issue salience, so that interest groups and politicians did little direct meddling and the influence of policy experts was high. Economic structuralism was consistent with traditional antitrust goals, and it also met the interests of institutional actors who reinforced each others' reliance on economics. For example, it offered the courts clearer decision rules than previously available. Linking market structure to firm conduct and performance, it lessened evidentiary burdens on prosecutors. Economic expertise enhanced antitrust agencies' administrative capacity and their autonomy from political actors. By increasing agency effectiveness, the ascendance of economics helped in responding to agency critics in Congress.

However, once economics dominated antitrust enforcement, change in its content reverberated through the antitrust agencies. When Chicago school economics replaced economic structuralism, the unintended consequence was to undermine the agencies' very purpose for being. With Reagan's election, evolutionary change from below, already producing "a Chicago-based enforcement agenda" (p. 215), coincided with revolutionary intent from above. In the 1980s, Chicago school economics came to *visibly* determine antitrust enforcement.

Within the political-institutional framework set, Eisner tells a thorough story, referencing interviews, and numerous primary and secondary sources. He provides excellent descriptive overviews of change in the law, policy, and institutions of antitrust and of the content of relevant

economic expertise. His argument that change in the latter has affected the former is persuasive. However, Eisner's book is limited by his failure to (1) provide a systematic, explicit set of hypotheses to be examined (or proposed), (2) situate antitrust in a comparative policy framework, and (3) place conclusions within a more complete theory of policy-making.

Without more systematic, explicit presentation of the full range of causal relationships suggested or implied, and data analysis tightly linked to these, the implications of Eisner's case study for general theoretical work on the relationships among law, science, and politics in policy-making may go undetected. Without comparative analysis and a *synthesis* of political-institutional and societal interest approaches to policy-making, Eisner's conclusions may mislead.

For example, Eisner argues correctly (p. 21) that political conflict conditions the emergence of professional expertise, but fails to note how societal, as well as political-institutional actors may struggle over technical expertise (see R. Stryker, "Science, Class, and the Welfare State," *American Journal of Sociology* 96:684–726). While his *theoretical* argument is that low issue salience explains societal actors' absence from the story of antitrust's transformation, his *history* indicates that intermittent aggressive enforcement against the interests of resource-rich regulated parties did cause these parties to promote change in antitrust policy and procedures (pp. 40–46, 60–62, 68–69, 174–79). Failure to analyze the impact on firms, industries, capital segments, and consumers of economics, and of the shift from economic structuralism to Chicago school economics neglects the politico-economic *functions* to expertise. It also misses possible feedback from societal actors to the political-institutional system.

In contrast, or in addition to Eisner's argument, perhaps Chicago school economics ascended and depoliticized in part because it pleased powerful regulated parties who bore the costs of economic structuralism. Perhaps not. We will know the conditions of autonomy, ascendance, politicization, or the eclipse of technical expertise in policy-making only by analyzing that expertise's impact on societal as well as on political-institutional actors, and by examining these actors' subsequent action or inaction vis-à-vis that expertise. We can have more confidence in our conclusions if we move toward comparative research designed to critically and explicitly evaluate plausible alternative explanations. Attention to the societal dimension, and to data from theoretically strategic comparisons would have added to Eisner's study, strengthening as well as conditioning some of its conclusions.

Rationalizing Justice: The Political Economy of Federal District Courts. By Wolf Heydebrand and Carroll Seron. Albany: State University of New York Press, 1991. Pp. xxii+308. $24.95.

Bryant G. Garth
American Bar Foundation

Rationalizing Justice is an ambitious, provocative, and, on the whole, convincing account of a major transformation in the federal courts. Wolf Heydebrand and Carroll Seron successfully integrate broad theoretical perspectives, historical sensitivity, and a structural analysis of inputs and outputs of the courts over time. So many fresh insights and interpretations emerge from the study that it would exhaust the space available even to list them. I will therefore focus on a few of the more provocative findings and some related questions of methodology and interpretation. In addition to making the review manageable, this approach should underscore the importance of coming to grips with and building on the insights of this work.

One key finding is what does and does not account for the dramatic federal caseload increase over the past few decades. Heydebrand and Seron demonstrate in several ways that, while population density contributes a decreasing amount of the explanation for increased filings, the role of the federal government—governmental presence—is becoming more and more important to caseload expansion (p. 65). They find indeed that "an explosion in public law litigation" (p. 176) has been transforming the courts away from their historical role as processors of disputes between private parties. This new role tends to give more power to the executive branch (e.g., p. 92) even in such matters as the resources and personnel available to particular courts.

A second notable finding is that the caseload increase has distinct effects on the output of the courts. Judicial productivity has improved in terms of dispositions per judge (p. 155), but there has been a marked increase in the number of cases resolved short of trial: "More and more cases are being decided at earlier and earlier stages under the supervision of an expanded team of court players" (p. 142). Different types of cases, moreover, are likely to get different treatments, in part because of the channeling of some cases to nonjudicial personnel (e.g., prisoner petitions that were initially handled by magistrates, p. 77).

The two general findings relate to a set of critical insights developed in the book. First, the authors find a "double bind" facing the federal courts. The double bind is that the courts face an increasing demand, coupled with lagging resources. The result inevitably limits the attention that the courts can pay to individual cases, again with results that apply generally and with specific effects on certain types of cases. Second, from a "due process" perspective (see p. 222) the character of the process is bound to change (p. 157). This kind of change moves the courts from a professional institution more toward what the authors term "the techno-

cratic model of justice" (p. 210). "Rationalized justice" applies then to the way in which it is dispensed and the rationalization that comes from accepting a model that departs notably from due process ideals.

The general picture is well-documented and supported by the empirical findings. Some questions come to mind, however, about the tone of the general conclusion as well as some of the details. There is a tendency in the book to employ crisis rhetoric: "The danger to the nature of the judiciary may not so much lie in bureaucratization, but in the potential shift in the role of the judiciary from a relatively autonomous branch of the federal government to an adjunct of the federal government, i.e., the executive branch" (p. 144); and "The crisis of the federal courts constitutes a social problem of major proportions with consequences that can be viewed as a crisis of the rule of law" (p. 194). It is difficult to distinguish crises from strains, as the authors recognize (p. 4). While I share the authors' concern, even alarm, with the developments they describe, I am skeptical about whether we face a real crisis. I doubt, in other words, if a continuation of present trends will have any great impact on the legitimacy of the courts, much less on the state in general (cf. pp. 203–4).

Beyond that question of interpretation and tone, a few specific questions ought to be raised. First, the rise of public law litigation (pp. 171–72) includes at least two components: one, emphasized by Abram Chayes, occurs when the "judge becomes the manager of public policy implementation," and the other occurs when "the judge becomes a low-level manager of administrative guidelines" (p. 200). We cannot tell from the reported data how much of the federal court work load falls in one category or another, but the difference is vitally important. We can see that civil rights cases and prisoners' cases tend not to get as much judicial attention as some other cases (pp. 147, 149–50), suggesting that judges may be capable of reducing this aspect of their work load through routinization and the use of other personnel. If the work load shrinks in that way, the rise of Chayes's version of public law litigation may be short-lived.

Private litigation may reassert itself. Indeed, the same ideology that supports a retreat from the liberal agenda favored by Chayes promotes the role of the courts in resolving business disputes (cf. pp. 72–73, 193, which suggest that business disputes will stay out of court; and consider the relative change from 1980–87 at p. 171). It may be that the federal courts are responding to strains (or crises) by reviving their historical interest in high-stakes commercial litigation and marginalizing the bulk of the public law litigation seen in prisoner petitions, individual civil rights cases, and social security appeals.

A related question concerns how to measure and assess a judicial change from adjudication to technocratic administration. The importance of the decline in trials, especially civil trials, is pointed out very clearly by the authors (e.g., p. 106). By 1988 only 5% of civil terminations were through trials. Unfortunately, the distinction between "no action"

and "pretrial" dispositions, which the authors use to highlight the pattern in civil cases, is not very helpful. The presentation of the data tends to suggest that pretrial means that the case was settled as a result of a pretrial consequence. This relationship probably has some viability after the amendments in 1982 to the federal rules of civil procedure, which then formally authorized judges to discuss settlements at pretrial conferences, but the reporting category used by the authors still applies only to cases in the time period that happened to be settled after a pretrial conference took place. Thus one cannot really say that "the procedural innovation [of the pretrial conference] began to take a larger share of civil dispositions [after 1940] . . . and stands at about the 50 percent level since 1970" (p. 118). The procedural innovation may have done nothing. The data that the authors report tell us only that pretrial conferences became more routine, and that therefore the normal tendency for cases to be settled on the eve of trial occurred more often at a time after which a pretrial conference happened to be held.

The more important point, however, which the authors well recognize, is that it will take more research to see exactly what is happening as the percentage of civil and criminal trials diminishes substantially (e.g., p. 157): "The specific effects of corporate and governmental forces on the nature and outcome of adjudication need to be studied by analyzing the qualitative process of decision-making itself, not just the quantitative statistical outcomes" (p. 218). Such analyses will allow a better understanding of what public law litigation is and is not, and where caseload pressure actually leads with more technocratic courts. *Rationalizing Justice* forces consideration of the critical qualitative issues, because it provides and analyzes the ample quantitative evidence which does suggest a decline in the quality and independence of the justice dispensed by the federal courts.

The Informal Economy: Studies in Advanced and Less Developed Countries. Edited by Alejandro Portes, Manuel Castells, and Lauren Benton. Baltimore: Johns Hopkins University Press, 1989. Pp. 327.

Gay W. Seidman
University of Wisconsin—Madison

For years, as an article in this collection says, the "informal sector"—broadly, the realm in which economic transactions go unregulated by the state—has been "a (pre)concept in search of a theory" (p. 78) and has included activities ranging from rag picking in Uruguay to electronics subcontracting in New York. The result of a five-year, four-continent collaborative project, *The Informal Economy* attempts to define and describe the dynamics of informality and to test common assumptions about why such relations seem to be increasing internationally.

At the global level, Castells and Portes suggest, the "expansion of the

informal economy is part of the process of economic restructuring" (p. 27) that has been going on since the 1970s. While capital seeks a cheaper, more flexible labor force, individuals and households try to survive a global economic crisis. Recognizing, however, that this broad analysis hardly illuminates the dynamics of the informal sector, the editors present a series of studies of unregulated activities in 14 cities, exploring through quantitative data and interviews with employers and workers the effects of the "process of informalization" on both individual workers' life chances and broader economic growth.

Several articles quantify and describe specific informal processes. Others analyze the close relationship between formal and informal production, looking at subcontracting and homework on the margins of regulated production processes. Finally, several articles discuss the role of states in creating and controlling informal sectors, either directly, by declaring certain activities illegal, or indirectly, by failing to enact or enforce labor laws.

If, as the editors suggest, new production technologies mean we can no longer expect to find industrial workers concentrated in large, easily regulated factories, what should we expect to find instead? The studies demonstrate why the informal economy has proved difficult to study or theorize. The nature of unregulated work is shaped by historically specific interactions between states, employers, and labor; work that is clandestine in one setting may be regulated in another. The details of how established industries may be built on informal labor are often fascinating, but the variation—especially between less-developed settings, where the alternative to unregulated work may be starvation, and more industrialized countries, where unemployed workers may still have access to social welfare programs—is enormous.

The empirical studies that make up the bulk of the volume test similar hypotheses about the impact of spreading informalization on regional economies. Does it offer new possibilities for economic growth? But generalizations, even systematic comparisons, prove elusive. In Italy's Emilia-Romagna, Vittorio Capecchi suggests, "flexible specialization"—small enterprises that can take advantage of new technologies—has allowed a new industrial growth strategy. But in other settings, unregulated work relations appear more exploitative; from Uruguay to Spain to Los Angeles, the descriptive studies conclude that expanded informality appears to be associated with cross-cutting and with stagnant or dependent technology, not innovation.

The contributors certainly prove the project's initial point: the informal sector is hardly an atavistic holdover. Unregulated production can expand through state prohibitions (as when the Soviet Union outlawed independent artisanal production), or selective legal enforcement by corrupt officials (as in Colombia's drug economy). But the editors conclude that spreading informalization in a context of international competition arises from "the efforts of firms and other private interests," often with

state acquiescence, "to gain market advantage by avoiding some state controls" (p. 299) such as minimum wage rules or taxes.

That observation underscores the importance of a historical understanding of informalization processes. As Patricia Fernandez-Kelly and Anna Garcia show in a comparison of Hispanic women workers in Los Angeles and Miami, and as Guy Standing's general discussion of Britain's informalization suggests, state policies draw the line between formal and informal work. Workers with household responsibilities may prefer working at home, but why should this work go unregulated? Especially where organized labor—centered almost by definition in the formal economy—has won state protection for regularly employed workers, spreading informalization is linked to changing relationships between states, employers, and workers. workers who are politically weakest—by virtue of gender, ethnicity, legal status, or skill levels—are least capable of inviting state control over wages, health and safety standards, and benefits.

If changing dynamics of state intervention shape the parameters of informality, the editors' policy recommendations for economic growth strategies seem somewhat paradoxical. State supports for informal enterprises are as likely to expand the formal economy—by increasing regulation of these enterprises—as to alter industrial growth patterns. Nevertheless, by detailing the varied dynamics of unregulated economies, this volume will frame future debates over what informalization is, and what its meaning is for regional growth.

Economics of Population Aging: The "Graying" of Australia, Japan, and the United States. By James H. Schulz, Allan Borowski, and William H. Crown. Westport, Conn. Auburn House, 1991. Pp. xiv + 364. $49.95.

Eileen M. Crimmins
University of Southern California

Much has been written in recent years about the disastrous futures in store for the countries of the world that are going to experience significant aging in the coming decades—much that is simplistic in argument, nonanalytic in content, and inflammatory in tone. This book is a valuable addition to the aging literature because it has none of these characteristics. It clearly and effectively explains why the economic impact of demographic aging is not a simplistic issue. The analyses presented lead to the conclusion that the doomsayers have been treading on shaky ground and that the impact of population aging is unlikely to be disastrous and is likely to be manageable for most countries without major sacrifices by the working population.

The authors' main theme might be stated "Demography Is Not Destiny." Most analyses of the future for aging countries rest simplistically

on the notion of changing old-age dependency ratios. This book provides a useful discussion of dependency ratios and a more sophisticated analysis than most other works of the meaning of the change in dependency ratios. The authors, unlike most other authors of works on aging, do not ignore the other dependent age group—the young—in considering the likely effects of the changing age structure of the future. This leads to a very different outlook for the effect of aging in the next 20–40 years. When the younger group is considered as well as the older dependent group, the "total dependency ratio will not be higher when the members of the baby boom retire than when they were children in the 1950's, 1960's, and 1970's" (p. 338). This relatively optimistic conclusion holds when the differential levels of both private and public expenditures for each of the dependent age groups are factored into the analysis.

The authors clarify that it is the rate of real economic growth in the future that will determine the ability of societies to provide support for their dependent populations without increasing the burden on workers. The factors that will determine economic growth are things like the level of "labor-force participation, savings behavior, investment in human and business capital, technological change, entrepreneurial initiatives, managerial skills, government provision of infrastructure, and so on" (p. 341). The authors conclude that aging of the population may have little to do with the rate of real economic growth. If economic growth is the solution to providing for our dependent population, the policies that will make it easier to adapt to an aging society will be those that will increase economic growth. Thus, the policy implications of this study contrast markedly with some of those that emphasize reducing expenditures for the dependent population.

While the volume is oriented around issues rather than the presentation of country studies, the comparison of Australia, Japan, and the United States is so rich in detail that virtually everyone, even those who are experts in the areas of retirement in one of the individual countries, will learn something new. These countries provide a very useful comparison group because of their economic and demographic similarities and their policy and programmatic differences. For instance, the Australian reliance on a means-tested noncontributory program to support the older population is unique among countries at this level of income. Other highly industrialized countries provide a universal social insurance system to support retirees similar to that in the United States. In addition, much detail is provided on the unique Japanese "two phased" retirement system, in which many employees take retirement from their lifetime jobs at age 55 long before the government system provides a retirement income. This system requires employers to play a more prominent role in providing for employees' retirement than is played by the private sector in other countries of similar economic and demographic development. Even those who are well versed in the details of the American social security and pension system may be unaware of the role played by macroeconomic conditions in the formulation and development of these sys-

tems. The details available on the institution of and the recent changes in the systems for providing retirement income in the three countries provide a useful source for anyone interested in one of the countries and a wealth for those interested in comparisons.

This book has many features to recommend it to a wide audience interested in aging, population, and economic/demographic/political interrelationships. It is extremely well written and easily understood by noneconomists in a wide variety of fields; yet because of its excellent review of relevant literature and original analyses, it is of value to economists and specialists in the areas of aging and retirement.

Fund Raising and Public Relations: A Critical Analysis. By Kathleen S. Kelly. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1991. Pp. xii + 539. $69.95.

Joseph J. Galaskiewicz
University of Minnesota

Kathleen Kelly presents a comprehensive volume on fund-raising within charitable organizations. Her purposes are to describe the function as it is currently practiced, offer an alternative theoretical perspective on it, and outline a set of strategies that practitioners can use to perform that function within the framework of this new theory. *Fund Raising and Public Relations* is written for practitioners and communications scholars, particularly those who train students in public relations. Nonetheless, students of organizational theory and philanthropic institutions should also find the book interesting.

Her main thesis is that fund-raising needs to be reconceptualized as a specialty of the public relations instead of the marketing function. This is necessary to counteract the growing influence of large donors over charitable organizations. Core programs, organizational missions, and operations are being unduly influenced by donors who are using their donations to leverage change in charitable organizations. The continued emphasis on fund-raising as a marketing function contributes to this problem.

Kelly rationalizes her recommendation by drawing heavily on organizational theory. It is well documented that organizations will redefine their goals and change their focus or activities to placate the demands and wishes of those providing the resources in their effort to ensure ready access to resources that can ensure their survival. Organizational autonomy is preferred, but the best an organization can hope for is to reduce the uncertainties surrounding resource procurement and negotiate terms that give it at least a modicum of autonomy. Kelly argues that this is the responsibility of those occupying boundary-spanning roles. In the case of charities, these are the fund-raisers and public relations specialists.

Problems arise within charitable organizations when the fund-raising function is only loosely coupled to the top management team and the board of directors. This happens when administrators and directors no longer involve themselves in this function on a day-to-day basis. When fund-raisers—all by themselves—confront the demands of big donors for increased control over the use of their funds, they find themselves torn between the organization's need for autonomy and the donor's wishes. The problem is exacerbated if fund-raisers are rewarded simply on the amounts raised each year. Then the incentives are to become the donor's agent *within* the organization.

If it is impossible to reestablish close ties between administration and fund-raising, then the only way to protect the organization is to change the way fund-raising efforts are evaluated. Evaluation needs to be based on public relations criteria, for example, the general condition of donor relations and effective two-way communication between the organization and its donors. Only then will fund-raisers come to see their job not as selling bits and pieces of the organization to the highest bidder but as boundary-spanning agents negotiating the degree of control donor publics will have over the organization.

To support her argument that donors threaten institutional autonomy, she cites aggregate statistics on gifts and giving, focusing in particular on the funding of higher education. She also provides detailed historical accounts of how foundations, corporations, and individual donors have used their gifts to leverage change. She augments these accounts with case studies on goal displacement within social service and cultural organizations. Kelly also draws heavily on her own personal experiences as a development officer in a large state university. To bolster her recommendation that fund-raising is a special branch of public relations, she retells the history of fund-raising and shows how changes in fund-raising practice paralleled changes in public relations practice. Although she never tested her "theories" or recommendations in any rigorous, scientific manner, she presents enough evidence to make her arguments credible.

Overall I applaud Kelly's bold efforts to conceptualize the fund-raising function within an organizations-environment framework and to redefine this function as a specialty within the field of public relations. Neither function is particularly well respected nor understood within the charitable organization, and her discussion helps to clarify the connection between the two. Kelly also attacks many commonly held assumptions in the fund-raising community. She shows that the ratio of restricted to unrestricted dollars is high, in-house staff often solicit donors, many organizations are dependent upon a small number of large donors (particularly in capital campaigns), and donors often attach strings to their gifts. She even dares to suggest that it is unethical for fund-raisers to manipulate donors, even if the action results in an unrestricted gift to the charity. In the fund-raising world, this book will be seen as "radical" and for this she is to be commended.

My only reservation is that the book does not say enough about donors' motivations or the microsocial worlds in which they live. The book is already over 500 pages long, and a summary of the literature on donors would necessitate another volume of equal size. Nonetheless, to outline a complete set of guidelines for fund-raising you need some theory about why individuals or institutions should give to the organization. If the donor is not going to have any more control over the organization, are there other incentives that fund-raisers can provide, for example, peer recognition, a sense of solidarity with the organization, and so forth? These issues are touched on, but they need to be discussed further.

Kelly is correct that fund-raising is an underresearched topic—especially in light of all the research on donative behavior. Searching out prospects, cultivating donors, and soliciting is probably as old as giving itself. Seldom does the latter occur without the former. If this richly descriptive and theoretically informed volume leads to more systematic research and theoretical work on this fascinating topic within the fields of organizational theory or communications, then I would deem it a success.

Corporate Society: Class, Property, and Contemporary Capitalism by John McDermott. Boulder, Colo.: Westview Press, 1991. Pp. xvi + 208. $44.00 (cloth); $16.95 (paper).

Maria Chan Morgan
Earlham College

John McDermott's ruling-class analysis of the mutual relationship between the modern corporation and society founders, like other class approaches, on methodological and ideological difficulties. Methodologically, he has not satisfactorily resolved "the empirical problem of determining the relationship of the dominant class to the political process" (p. 107), in spite of his claim to the contrary. Ideologically, his commitment to egalitarian social democracy leads him to see new conspiracies even as he acknowledges the untenable positions of the old "conspiracy theories" (his words) of C. Wright Mills, G. William Domhoff, and others.

McDermott argues that the basis of social and political domination has shifted historically. In the past 100 years, changes in technology have given rise to a modern corporate form, in which the top management and the middle "multicompetent" professional classes have almost total control of and access to corporate property, subjugating the modern working class to what he calls social Taylorism—"the phenomenon of socially compulsory consumption, and the transformation of trade unions . . . into controlling arms of the state mechanism." (p. 145). Yet in conclusion, he concedes that workers are better off in modern corporate society than at any other time.

Painfully aware that his argument sounds very similar to the largely discredited power elite theory of C. Wright Mills, McDermott insists that his focus on the property structure and the middle professional class of the corporate form differentiates his analysis from other elite or class approaches. McDermott argues that the modern corporation is more than just a Weberian bureaucracy, it is a "class structure of production" (p. 4) or what he calls, in the rest of his book, a corporate form. While a cohesive coalition of the top and middle classes will resist most of the demands made by workers, they will concede some social rights to the working class as long as they are compatible with increasing the productivity and profit of the corporation. This produces what McDermott calls the great paradox of corporate society: on the one hand, "workers must submit to the technical-productive discipline of modern capitalism," on the other hand, "they are given limited rights in the enterprise and extensive rights in the society" (p. 190).

To McDermott, class cannot be separated from property relation (p. 6). Just as class is a historical concept, so is property. Instead of being largely private family property, the modern corporation is quasi-collective corporate property, with top management as a group exercising the classic rights of property over the physical, commercial, and above all, human assets of the company. How such a definition of collective property can be independently verified is left to the reader. To compound the methodological problem, McDermott adds individual and social *behavior* in addition to property as another criterion of class definition.

Although he suggests earlier in the book that the three classes in the corporate form have *both* joint and conflicting interests (p. 5), the bulk of his book focuses almost entirely on class antagonism. As might be expected of a ruling-class analyst, McDermott rails against the usual left-wing targets: the military, big corporations, universities, government—especially defense and national security agencies—"companies" (according to his definition organizations like the Rand Corporation and the Hoover Institution), Commerce Department statistics, mainstream social science, and equal opportunity. Of course, the top-management class is criticized for being responsive only to the corporate public consisting of institutional investors, bankers, auditors, accountants, lawyers, and government regulators. The middle self-supervising professional class is attacked for exercising "industrial dictatorship" (p. 188) over the working class.

Given the standard ruling-class critique of modern government and business relations in the bulk of his book, McDermott's conclusion about the improved position of the modern working class comes as a surprise. He notes that the social power and the political influence of the working class have been strengthened and expanded in corporate society. But instead of vindicating pluralist capitalism, McDermott gives credit only to the "adaptive, innovative, and autonomous intelligence" of the workers for refusing to succumb to the "cultural hegemony of the dominant classes" (p. 191). If this indeed is true, McDermott owes it to the workers

to devote more space in his book to their inventiveness and less to the adaptive ingenuity of creative industrialists like Carnegie and Sloan.

Organizational Ecology. By Michael Hannan and John H. Freeman. Cambridge, Mass.: Harvard University Press, 1989. Pp. xvi + 366. $29.95.

Ronald S. Burt
Columbia University

Organizational Ecology is an exceptional book. It is at several points an exemplar of integrating formal theory with sophisticated empirical research. It is certainly the finest sustained substantive argument based on the statistical methods that Michael Hannan (with Nancy Tuma) introduced in the 1970s. The integration of theory with data analysis is strongest regarding the liability of newness, mixed regarding density dependence (where competition and legitimacy are invoked), and weakest regarding the environmental variables used to measure contextual factors affecting population growth. The book is the foundational statement of what has become an important component in the sociology of organizations: population ecology analysis. But there is interest beyond the organizations specialty. The book is instructive for anyone analyzing competitive arenas. In a graduate seminar, I found it to be a powerful teaching device because of the conflict it generates, forcing students to justify their beliefs about organization and competition.

The authors' premise is that interesting things about organizations are revealed by an analogy with populations of biological organisms. In particular, biological models of how populations survive and prosper on the resources of a specific ecosystem offer insights into the survival and growth of organization populations.

Organizational Ecology is a general statement of this premise, along with empirical evidence. It brings together basic lines in the authors' work over the past decade (with helpful ties to work by close colleagues such as Glenn Carroll). The first four chapters lay the foundation. The authors review the sociology of organization literature to position the population ecology approach clearly, then they introduce basic concepts (organization, population, structural inertia, selection vs. adaptation). Chapters 5 and 6 are the key theoretical chapters that define competition within a niche, niche-carrying capacity, and density dependence. Chapters 7 and 8 review statistical methods and describe the organizational populations used for evidence in the book. Results are presented on the populations of national labor unions, semiconductor producers, San Francisco newspapers, and California restaurants. Chapters 9–12 provide empirical illustration of the hypotheses. Chapter 13 highlights some interesting directions for subsequent research.

The strong link between theory and evidence makes it possible to characterize the book by its strongest empirical results. I find the results on the liability of newness the most powerful in the book. The idea is that advantage goes to the organization with reliable, accountable, and reproducible behavioral scripts. These rewarded qualities jointly create structural inertia in organizations, which sets the stage for population ecology's emphasis on selection processes over adaptation processes. For various reasons, new organizations are less skilled, and less perceived by suppliers and consumers as skilled, in their reliability, accountability, and reproducibility. The liability to new organizations dissolves with time, according to an unknown function that will have to be recovered empirically. Ample evidence is presented to support the liability of newness and the evidence is nicely suited to the methods and data.

The density effects are equally clear empirically but more ambiguous in theory. The number of organizations in a population (N) is discussed as density. As the density of organizations feeding on the same resources increases, competition intensifies and growth slows (fewer births, more deaths). At the same time, the legitimacy of the known forms of doing business in the population is established, which dampens mortality. The combined result is that growth is especially fast at low density and slowed at high density (e.g., negative effect of N on mortality and positive effect of N^2). However, the specific conditions that generate competition and legitimacy are ambiguous (or, more accurately, defined by their outcomes rather than their substance).

This can be clarified by referring to the very different analytical stances now associated with familiar market approaches as opposed to those of population ecology. Market approaches look for structural and behavioral conditions that generate competitive advantage. The population-ecology approach infers competitive advantage from relative growth rates. This seems a natural circumstance for some powerful collaboration between the alternatives inasmuch as each is ignorant where the other is strong. Population ecology stands to gain powerful models and hypotheses about constrained access to environment resources to use as growth covariates in comparisons across populations, while the market approaches would gain powerful models and hypotheses about markets over time.

There are also important questions raised. Given the clear analogy between market and niche, biological models of stable networks of interdependent populations within an ecosystem raise interesting questions (and answers) about stability in interorganizational networks. At the same time, the clear analogy between market and niche raises a question of whether firms or establishments should be the organizational unit counted in population-ecology models. Density, N, is now a count of firms. Should this a count be of establishments? The establishment is the organization unit affected by competitive forces in the market approaches and also in the population-ecology approach. The establishment better corresponds to the single biological organism with known resource needs

(in the organization case defined by technology). This shift might have little effect on the restaurant or newspaper populations studied in *Organizational Ecology* (ignoring the obvious complications of corporate chains and franchising), but has implications for populations in the stronghold of the market approaches—manufacturing where large firms typically operate many establishments in several markets (i.e., in several population niches). The beginnings of this can be seen in the strong results reported in *Organizational Ecology* when ecological models are applied to independent semiconductor firms versus the poor results obtained when the models are applied to plants owned by larger firms. Work on these issues will no doubt appear over the next few years as population-ecology insights are integrated with more established approaches.

The future will look back to the late 1970s through the 1980s as the period when theory emerged to formalize the temporal component in the relationship between competition and formal organization. The key book in that long look back will be Hannan and Freeman's *Organizational Ecology*. This is a book in which we can all take pride as sociologists.

The AIDS Disaster: The Failure of Organizations in New York and the Nation. By Charles Perrow and Mauro F. Guillen. New Haven, Conn.: Yale University Press, 1991. Pp. xii + 206. $25.00 (cloth); $9.95 (paper).

Peter Bearman
University of North Carolina at Chapel Hill

Ten years have passed since AIDS was first recognized by the Centers for Disease Control. In this short period, 160,000 Americans have been diagnosed with AIDS, and 110,000 have already died. At current rates, 30 people die from AIDS every day, and, every 15 minutes, somebody is diagnosed as having AIDS. Somewhere between 1.5 and 2.5 million Americans now carry the HIV virus that leads to AIDS. The AIDS cases we see now are the result of risks people took five or 10 years ago. The AIDS cases we will see in the year 2000 will be a result of risks people have taken (i.e., having unprotected sex and sharing needles) over the past couple of years. These are the basic facts, and it is in this context that Charles Perrow and Mauro F. Guillen have written *The AIDS Disaster*.

The transmission of the HIV virus can be prevented simply—by using latex condoms during sex and by using clean needles for the injection of intravenous (IV) drugs. And as the authors note, condoms and bleach (to sterilize needles) are available in most drug stores at marginal cost. Thus the fact that the HIV virus continues to spread reflects a profound failure—the failure of organizations responsible for education and for outreach, but also the failure of individuals. Since it is reasonable to believe that if organizations had responded quickly and appropriately to AIDS in the first years of the epidemic (1981–84), the crisis we confront today would be significantly less severe, Perrow and Guillen argue that

we must understand why organizations failed in order to respond to the challenges of AIDS today.

There is sufficient evidence to show that organizations failed to deal with AIDS. For three years, bloodbanks refused to take the steps necessary to protect the blood supply. As a result, 95% of factor VIII hemophiliacs are now HIV positive. Later, the blood industry stalled on look-back programs, thereby preventing recipients of HIV-contaminated blood from learning of their potential HIV status. The government failed as well—across all levels—from the cities to the federal government. The history of the response to AIDS is not pretty. Part of this book presents this story and Perrow and Guillen do a good job here, but more compelling is Randy Shilts's *And the Band Played On* (New York: St. Martin's, 1987).

Where Shilts is weakest—making sense of organizational failure beyond individuals—Perrow and Guillen are strong. Their view is that AIDS presents unique and unusually difficult problems for organizations that in the best of times are prone to failure. These problems, and they constitute quite a list, are (1) that AIDS is seen as self-induced by people who engage in illegal risk behaviors; (2) that AIDS hits the hardest in the outcast communities, especially among the urban poor, that is, minorities, homosexuals, intravenous drug users, and prostitutes; (3) that AIDS is frightening to the public and that it appeared at the same time as a powerful conservative movement, whose leaders felt more comfortable pursuing as a basic goal the "conservative fantasy of how people ought to behave" (p. 125) than policies that could have stemmed HIV transmission; and finally, (4) that AIDS interacts with and exacerbates other social problems: the crisis of our health system, intravenous drug abuse, the diffusion of other sexually transmitted diseases, homelessness, and the hopelessness of the young in the urban ghettos. Perrow and Guillen argue that the failure of organizations responding to AIDS cannot be understood outside this context.

It is in the poor communities of the major urban cities that HIV/AIDS is now spreading most rapidly, in association with intravenous drug use and the crack epidemic. It need not have gone this route, but the transmission system of AIDS now means that we can, in the absence of effective policies to stem disease spread, expect infection to saturate the high-risk populations and then to trickle down to the lower-risk groups. Local organizations, especially in the minority and intravenous drug use (IVDU) communities, are especially critical for outreach and health care. Few organizations are in the position to be effective. Perrow and Guillen track a sample of local organizations during the AIDS epidemic and discuss cogently their strengths and weaknesses. The basic message is that few coped with AIDS very well. Not a trivial number of them simply broke down in response.

Perrow and Guillen identify five organizational responses to AIDS: (1) the incorporating of AIDS programs, with a resultant enrichment of other programs, (2) denial and flight (federal government, NYC housing

authority, (3) the segregation of HIV/AIDS programs (the response of the state and local administration in New York), (4) goal distortion (schools, Catholic Archdiocese shelter), and breakdown (hospital special teams, families). There is no unitary finding.

Drawing from in-depth interviews with agency heads and members and representatives of organizations (many of whom were openly hostile to the authors), Perrow and Guillen uncover the jagged and irregular sides of organizations struggling to define goals and strategies in a context defined as much by ideological opposition and fiscal constraint as by tangible need. Few succeeded. Centralized and bureaucratic organizations failed as profoundly as decentralized organizations staffed by volunteers. Local units failed, as did larger state and federal units. Each failure generated additional problems for other organizations—hospital bed shortages created problems for shelters for people with AIDS, inadequate funding of IVDU treatment centers meant that men and women seeking treatment for addiction would have to wait months for a vacancy, leading to further HIV transmission—and so on. Perrow and Guillen provide flavor to the ecology of AIDS in a population of related social problems. Their analysis leads them to call for a massive intervention program along the lines of the Polaris missile program. Even Desert Storm would be an improvement.

This book has weaknesses. Readers may find the tone too normative for their tastes. The diffusion model which underlies the text is simplistic, neglecting the pathbreaking work of modelers from Los Alamos (see J. Hyman and E. A. Stanley, *Proceedings of Mathematical Approaches to Environmental and Ecological Problem Solving* 81:190–219) or Michigan (see J. Koopman et al., *Journal of AIDS* 1:486–504) whose work on mixing models appears basic to our understanding of AIDS epidemiology. The organizational analysis is not as crisp as one might expect, especially given the standards set by Perrow in *Complex Organizations* (New York: Random House, 1986), and the in-depth interviews that provided much of the documentary evidence seem underdeveloped in the text. I kept looking for more depth. These weaknesses aside, *The AIDS Disaster* does not fail. This book will shape sociological models of organizational response to the AIDS epidemic in the years to come. This disease is important, even if we would like to think it away. Perrow and Guillen have provided us with lots to think about.

Rhetoric in an Organizational Society: Managing Multiple Identities. By George Cheney. Columbia: University of South Carolina Press, 1991. Pp. xi + 202. $29.95.

Gene Burns
Princeton University

In *Rhetoric in an Organizational Society,* George Cheney draws on the writings of Kenneth Burke to analyze the rhetorical process by which organizations establish a collective identity that transcends individuals. In doing so, organizations must manage multiple identities to the extent that they have a diverse membership and address disparate audiences. The book is written from the perspective of the field of communications and emphasizes the internal structure of rhetoric, but it places itself theoretically in a literature that has influenced a variety of social scientists. It merits examination by sociologists with empirical interests in religion, organizations, and perhaps political culture A number of potential readers may miss the volume, however, because the title gives no clue that the book is primarily a case study of the National Conference of Catholic Bishops' 1983 pastoral letter on peace and nuclear strategy, *The Challenge of Peace* (Washington, D.C.: United States Catholic Conference, 1983).

Indeed, *Rhetoric in an Organizational Society* is the most informative social scientific study to date of the American bishops' recent forays into public policy debates. Cheney's perspective is the source of a number of important insights. For example, he emphasizes the difficulty bishops had in applying their religious authority to a secular policy debate. They sought to influence such debates without being identified too strongly with any particular political faction. This serves as an excellent empirical illustration of the "multiple identities" that are Cheney's theoretical focus.

In addition, the author points to the bishops' use of ambiguous statements to avoid some specific, controversial commitments. For example, the bishops wanted to increase the moral weight of pacifism but without seeming to devalue Catholic just war doctrine. To avoid the appearance of contradiction, the bishops glossed over the differences between the alternative approaches and made the questionable assertion that the two are independent because of a common presumption against the use of force. Cheney's emphasis on the sources of rhetoric also results in a sophisticated understanding of some of the constraints and opportunities the bishops faced. He chronicles well the doctrinal developments within international Catholicism that contributed to *The Challenge of Peace.* And he notes that the Catholic church, like any organization or person, could not safely assume the posture of critic until it had become an insider. Thus, the bishops' critical view of the U.S. government's policy of nuclear deterrence could not have emerged before the church's ascen-

sion into mainstream American society, symbolized by the 1960 presidential election.

However, the book's perspective is also an obstacle, both to the reader and to the argument. The focus on the rhetorical devices present in the pastoral letter leads occasionally to tedious, confusing, and repetitive discussions; it also results in an unclear presentation of the substance of the bishops' argument. Furthermore, a number of central historical and institutional factors are overlooked. One would never know from the book that Vatican II was a watershed in Catholic history or that the U.S. bishops' identity as Americans has led them to diverge from European Catholic views of politics for over a century. That the letter suggests the possibility of new types of church-state conflict, and may represent a bold attempt by American bishops to expand their influence within the international Catholic church, does not receive much attention.

The overemphasis on rhetoric often leads to the functionalist assumption that organizations necessarily accommodate their multiple audiences. The audiences themselves are taken for granted, of interest only to the extent that they affect rhetorical choices. This leads to a number of methodological and substantive problems. First, the book only fleetingly and anecdotally mentions how Catholics and Americans actually responded to the letter. But the management of multiple identities is fundamentally an attempt to satisfy and manipulate diverse audiences (p. 6); we need to look at those audiences in order to evaluate the bishops' success. Second, Cheney's functionalist assumptions force him to conclude (p. 168) that *The Challenge of Peace* was an inherently conservative project. On a number of issues the bishops certainly backed away from commitments that could strongly challenge Washington or Roman doctrine But it is also true that the bishops knowingly invited controversy and, within Catholic thought, were quite bold and innovative. Part of the boldness lay in choosing to *expand* their audience. A more conservative approach, taken by many popes and bishops in the past, would have been to continue to address only a narrow, more homogeneous constituency. It was a daring undertaking to enter contemporary public policy debates with a detailed critique of prevailing assumptions. Instead of being a short, moralistic pronouncement, the 1983 letter was a small book full of specific policy recommendations.

Despite these problems, *Rhetoric in an Organizational Society* is essential to an understanding of the new paths blazed by American Catholic bishops.

He-Said-She-Said: Talk as Social Organization among Black Children.
By Marjorie Harness Goodwin. Bloomington: Indiana University Press,
1990. Pp. x + 371. $49.95 (cloth); $19.95 (paper).

William A. Corsaro
Indiana University

"Intellectual breadth shines through this book," notes Barrie Thorne in
a promotional blurb that is, for this impressive monograph on talk as
social organization, not in the least exaggerated. In fact, Goodwin's care-
ful and insightful scholarship is one of many scholarly traits that display
the influence of one of her mentors, Erving Goffman. Like much of
Goffman's work, this monograph is full of detailed discussion in the text
and footnotes that link data analysis and conceptualization to the relevant
literature. These discussions literally "confront" the work of others and,
in the process, immerse the reader in Goodwin's data and her quest to
document how her analyses question, refine, and expand the best work
in the field.

A second quality Goodwin's work shares with that of Goffman is a
profound concern with "units" of social action. Although Goodwin spec-
ifies local *activities* as the basic unit of analysis (pp. 8–9) in her research,
she, like Goffman, constantly struggles with the perplexing fact that any
strip of social activity is always embedded in, while at the same time
capable of transforming, some larger unit. Thus, Goodwin introduces
(usually without definition) over 20 different unit terms or descriptions.
In nearly every case these terms are analytically useful and demonstrate
both the complex nature of social discourse as well as the necessity for
integrating conversational analysis with in-depth ethnography.

Goodwin's analysis of social interaction moves beyond the work of
Goffman in one important respect. While Goffman often relied on sec-
ondary sources or hypothetical examples, Goodwin's work is based on
detailed analysis of audio-taped, naturally occurring peer conversations
that were collected as part of an intensive ethnography of urban, black
working-class children in Philadelphia. Blending analytic techniques bor-
rowed from conversational analysis and Goffman with ethnographic in-
formation on the community, the children, and the peer culture, Good-
win, through the investigation of the form and sequencing of verbal
activities, demonstrates how the social order of the moment is formu-
lated, rejected, and reconstituted through talk.

In the first three chapters Goodwin lays out her theoretical perspective
and describes her fieldwork methods, the ethnographic context, and the
participants. The theoretical chapter starts well, but it becomes some-
what disorganized and ponderous. After a crisp introduction in which
Goodwin argues that her book "is about how talk is used to build social
organization within face-to-face interaction," and that such interaction
"constitutes a central place where members of a society collaboratively
establish how relevant events are to be interpreted, and moreover use

such displays of meaningfulness as a constitutive feature of the activities in which they engage" (pp. 1–2), she then locates her perspective and the later analysis in related theoretical and empirical work. Here Goodwin's discussion becomes more uneven. She presents an informative review of conversational analysis. However, she says little about Goffman and other students of discourse analysis on whom she relies in the later analysis nor does she address some of the basic differences between conversational analysis and other approaches to the study of discourse. Chapters 2 and 3 are well-written, insightful descriptions of the method and setting. Chapter 3 depicts (with descriptions and photographs) the setting (a black working-class neighborhood in Southwest Philadelphia) and participants (15 girls and 23 boys aged 9–14), describes the children's culture, and integrates the study with related work on children's peer interaction and culture.

The eight analytic chapters in the book each address key "activities" in the children's culture: directives (talk designed to get someone else to do something), argument, "he-said-she-said" (a girls' gossip-dispute process), instigating (a way of promoting he-said-she-said confrontations), and stories (p. 9). These chapters are astonishing in both the sheer empirical detail of her documentation of the production of the particular activities and the more general theoretical implications she draws (most especially those related to gender differences in the way children organize their activities). I have space for only one example of each of these qualities. Goodwin's analysis of a boys' task activity demonstrates how directives are best seen as actions embedded within a larger field of social activity. Thus, they have "a Janus-like quality, on the one hand looking outward as they link what is happening at the moment to larger frameworks of meaning and action, and on the other looking inward as they invoke coherent scenes, self-sufficient worlds within which images of participants are animated" (p. 108). Regarding gender differences, Goodwin's findings challenge commonly accepted beliefs that boys are more competitive and girls more cooperative. Her analysis of girls' disputes are especially instructive. Unlike the direct competitive disputes of males, black females frequently engage in gossip disputes during which absent parties are evaluated. The airing of grievances frequently culminates in he-said-she-said confrontations in which one girl, A, challenges another girl, B, about what B told a third girl (C) about A. Goodwin's analysis specifies how "girls can call forth feelings of righteous indignation that are relevant to a dispute which can last over a month. The interactive frame created by the he-said-she-said event thus links together several speech events—stories, gossip, and argument—within a single process, the management of breaches" (p. 286).

Goodwin's conclusion, like her introduction is brief. She touches on but fails to develop the more general implications of her work for (1) the debate between conversation and discourse analysts, (2) differences between cognitive and interpretive anthropologists, and (3) collective versus individualist approaches to childhood socialization.

This hesitancy to develop more fully the general theoretical implications of her work is, however, a minor shortcoming. Overall, this is an excellent monograph that can serve as a model of quality scholarship in advanced undergraduate and graduate courses in social psychology, socialization, sociolinguistics, and qualitative methods. Although her work is highly respected among sociolinguists and an increasing number of anthropologists and sociologists, Goodwin's research is not as well known as it should be. The book brings together and extends Goodwin's best work, and its publication should bring her work the attention it deserves from a wider audience of social scientists.

Ideology and Modern Culture. By John Thompson. Cambridge: Polity Press; Stanford, Calif.: Stanford University Press, 1990. Pp. vii + 362. $39.50 (cloth); $12.95 (paper).

Douglas Kellner
University of Texas

John Thompson argues in *Ideology and Modern Culture* that ideology in the modern world is intrinsically bound up with mass communication and that therefore an adequate analysis of ideology requires analysis of its transmission by mass media of communication and its reception by mass audiences. In carrying out this project, Thompson provides analyses of the concept of ideology, the role of ideology in modern societies, the concept of culture, cultural transmission, and mass communication, the connections between ideology and mass communication, and an interpretive methodology for culture and social theory. In each case, Thompson provides a clear account of the history of the concepts at stake, makes salient conceptual distinctions, and provides a broad social theoretical account of the phenomena under investigation. His studies should therefore prove of interest to a broad range of social theorists and those interested in mass communication, ideology, and the constitution of modern societies.

In an earlier work, *Studies in the Theory of Ideology* (Berkeley: University of California Press, 1984), Thompson developed a theory of ideology as a form of discourse. Building on theorists like Castoriadis, Lefort, Bourdieu, Giddens, Gouldner, Faye, Pêcheux, and Habermas, Thompson analyzed ideology in terms of language and power and the ways that ideological discourse functions in the social world. In *Ideology and Modern Culture,* by contrast, he interprets ideology as a symbolic form connected with the symbolic transmissions of mass communication. In both books, however, he is concerned to develop a critical concept of ideology that defines ideology as discourse and symbolic forms that sustain relations of domination, legitimating relations that involve asymmetrical relations of power (pp. 7, 151). He stresses that ideology does not merely involve class domination, as in the classical Marxian concept, but a wide variety of forms of gender, racial, and national forms of domina-

tion. Ideology for Thompson thus involves the mobilization of meaning in symbolic forms to sustain domination and to serve the interests of ruling elites.

After preliminary discussions of his program and strategy, Thompson begins by reconstructing the history of the concept of ideology. In Thompson's narrative, ideology had a primarily positive sense in the work of Destutt de Tracy and his colleagues in the Institut National who, after the French Revolution, pursued the Enlightenment project of tracing the origins of ideas in sensations and experience; this genealogical study of ideas was called "ideology" and was conceived as a "first science" upon which a foundation of knowledge could be built. For Napolean, these "ideologues" were far removed from practical affairs and reality, and he thus provided the concept with a pejorative resonance—a position taken over and developed systematically by Marx. Thompson suggests that Marx has several different conceptions of ideology (pp. 33 ff.), though he generally interprets ideology as an instrument of class domination.

Later social theorists like Mannheim developed a more neutral, sociological, explication of ideology simply as dominant ideas shared by major groups, including the working class. Building on the critical concept, Thompson wishes to keep the connection between ideology and domination, but construes denomination more broadly than do orthodox Marxists. He also criticizes what he calls the "grand narratives" of modern ideology, developed by Marx and Weber, for omitting the role of mass communication in the transmission of ideology. In addition, he criticizes the general theory of state-organized capitalism of Althusser and his school for similarly omitting mass communication and developing too monolithic a theory of ideology. While the Frankfurt school connects ideology with mass communication, they, too, are overly monolithic and reductive in their analyses of ideology.

To overcome the systematic deficiencies of previous theories, Thompson grounds his concept of ideology in theories of culture, mass communication, and hermeneutical social criticism. He provides excellent historical discussions of the themes under consideration and indicates the relevance of his discussions for contemporary social theory. If I had to venture a critique of Thompson's enterprise, it would be that he attempts too much in one book when he tries to develop (1) a theory of ideology in modern societies; (2) critical theories of mass communication and contemporary society; and (3) a method of social and cultural interpretation and critique. Separate books could be written on each of these topics, and Thompson promises further elaboration of his conceptual perspectives in future work.

There is one major objection to the concept of ideology, however, that Thompson does not confront. Postmodern theorists like Baudrillard, Jameson, and others have argued that there is an implosion between ideology and culture in contemporary societies, such that it is impossible to distinguish between ideology and nonideology when all mass commu-

nication and public discourse is ideological. They also argue that mass communication is so fragmented and superficial that one can no longer find a depth of meaning in its images, figures, and discourse of the sort that Thompson wants to explicate in his method of "depth interpretation." I do not myself buy the postmodern critique (see Steven Best and Douglas Kellner, *Postmodern Theory: Critical Interrogations,* London and New York: Macmillan and Guilford Press, 1991), but believe that Thompson's position would be stronger if he would engage these postmodern positions and provide a rebuttal to their more extreme claims— which put social theory itself into radical question. Thompson restricts his focus, however, to ideology and *modern culture,* thus failing to engage postmodern discourse and critique. Yet, as an explication of the vicissitudes of ideology in modern theory, Thompson's work is a tour de force full of provocative ideas, sharp analysis, and constructive social theorizing.

The Conscience of the Eye: The Design and Social Life of Cities. By Richard Sennett. New York: Alfred A. Knopf, 1990. Pp. xiv+266. $24.95.

Daphne Spain
University of Virginia

In his introduction to *The Conscience of the Eye,* Richard Sennett says that the book is the third part of a trilogy, the first two parts being *The Fall of Public Man* (1977) and his novel *Palais Royale* (1986). In the current volume he attempts to combine the social criticism of *Public Man* with the personal insight of *Palais* to relate "architecture, urban planning, public sculpture, and the visual scenes of the city to its cultural life" (p. xiv). It is possible to see the origins of this book even earlier than 1977, however. Its themes closely resemble those presented by Sennett in 1970 in *The Uses of Disorder.* In this earlier work, Sennett proposed that people avoid those different from themselves in an attempt to create "purified communities." The current organization of the city encourages the affluent to escape disorder by moving to homogeneous suburbs. Sennett thinks that certain kinds of disorder should be encouraged so that people's tolerance for diversity increases; as the exposure to confrontations expands, the tendency toward violence declines. The ability to accept and live in disorder represents the maturation of a community's citizens.

The Conscience of the Eye applies similar ideas more concretely to urban form by asking how this aversion to "otherness" shapes our architecture and cities. Its central theme is that current urban design reflects a collective fear of exposure to the unknown, an "unknown" that contains more threatening experiences than stimulating ones. Flight from the unknown results in sharp divisions between the interior and exterior

spaces of daily life. Walls of plate glass separate inside from outside, for example, while highways bisect once thriving neighborhoods. Such barriers serve to neutralize places of public contact and further highlight differences among city dwellers. The challenge for Sennett is how to "revive the reality of the outside as a dimension of human experience" (p. xiii).

Sennett's attempts to find answers to this challenge take him far back and far afield. He first targets the Judeo-Christian ethos for its emphasis on the fears of spiritual dislocation and exposure, then turns his lens on Puritanism's need for refuting worldly distractions to find peace within oneself. This religious zeal to reject the exterior world in order to master the interior one has resulted in modern urbanists who are "in the grip of the Protestant ethic of space" (p. 42). By this Sennett means that urban planners (and one could add architects and real estate developers) tend to neutralize environments out of a fear of pleasure; otherwise how can one explain the sterility of most shopping malls and public plazas?

Moving forward in history, Sennett finds a metaphor for our modern urban dilemma in the *émigration intérieure* of Louis Philippe's reign. This was the term applied to the aristocracy's voluntary withdrawal from public life as a protest against the king's corrupt materialism. By retreating to their country estates, these noblemen made a clear gesture of their need for sanctuary from the secular world. This search for sanctuary and its spatial manifestations is a continuing theme throughout the book. One of the products of the *émigration intérieure* was Tocqueville, whose journey to America impressed him with the blandness of its cities. Sennett attributes Tocqueville's reaction to the already extensive use of the grid—a pattern Sennett calls a "weapon to be used against [the] environmental character" of cities (p. 52).

This combination of Judeo-Christianity, Puritanism, and Tocqueville is but the first of many strands woven by Sennett into his tale of urban design. His tapestry includes reference to architects, planners, philosophers, and novelists. He is more likely to cite Palladio, Mies van der Rohe, and Le Corbusier than Robert Venturi and Frank Gehry, however, and more likely to cite Baron Haussmann and John Wood the Elder than Ed Logue or Andres Duany. Among philosophers whose work Sennett applies to the urban condition are Nietzsche and Hannah Arendt; the novelists include Proust, Flaubert, and James Baldwin.

It is from James Baldwin that Sennett takes his lessons about the necessity for the "narrative space" of cities Cities should be designed to reflect the intersection of space and time represented in novels; that is, the overlapping of events and people then create a rich and complex story. This narrative space can be created by changing the way open spaces and buildings are conceived, which means changing the nature of spatial and temporal boundaries. Planners and architects creating places with narrative power would make them multifunctional for different uses throughout the day (an endorsement of complexity made by Jane Jacobs over 30 years ago).

It is toward the end of this exegesis that we begin to suspect the "D" word is lurking. Indeed, on page 224, Sennett finally names his image as "a city of deconstructions." Yet he denies that he is proposing a typically deconstructionist approach to differences, discontinuities, and disorientation. Rather, he sees these forces as ways to expose people to one another through ethics of sympathy and empathy. The urban conscience of the eye is the power to arouse empathic concern for disorientation and differences among people. Only then can individuals center themselves in relation to others.

This language may seem somewhat vague because it is a reflection of the book's emphasis on conceptual and philosophical issues to the almost total exclusion of policy. Therefore the book may be frustrating to planning practitioners. For sociologists and urban theorists, however (as well as certain segments of the general public), this should be a thought-provoking addition to our ways of thinking about city form. Sennett continually urges us to reject simple treatments of the city in order to more fully appreciate the uses of disorder.

The Cynical Society: The Culture of Politics and the Politics of Culture. By Jeffrey C. Goldfarb. Chicago: University of Chicago Press, 1991. Pp. xi + 200. $22.50.

Richard Flacks
University of California, Santa Barbara

Widespread public cynicism, Jeffrey Goldfarb believes, is the single greatest barrier to democratic possibility in contemporary America. Cynical attitudes undermine democratic debate because both opinion makers and ordinary citizens have come to see policy proposals merely as efforts to engineer mass support instead of examining their substantive value. All actors who claim to be serving the public interest are seen as really self-serving. A mocking cynicism is directed at critics of the status quo—destroying their pretensions to integrity and thereby reinforcing a pervasive mood of political resignation. Established forums of public debate—such as party conventions and legislative bodies—are universally regarded as nothing but arenas for political dogfighting. Educational institutions are increasingly organized as mechanisms of "quality control" rather than as frameworks for individual and cultural development. Democracy requires the sense that people are speaking authentically, that ideas matter, that public dissent be taken seriously—and that citizens believe in the possibility of a common good and a common ground that transcends private interest. In a cynical society, such beliefs are hardly expressed—for to do so is to open oneself up to ridicule and suspicion.

Goldfarb, therefore, has undertaken here a rather courageous intellectual project, and the results deserve close attention. First of all, to have forced notice of cynicism and its political and cultural symptoms is, in

and of itself, a signal contribution. But Goldfarb goes beyond insightful diagnosis to sketch an analysis of the social and cultural roots of cynicism. To do this, he resurrects the concept of "mass society," arguing that contemporary American political culture can only be understood as an uneasy combination of democratic and "mass" modes of integration. The tension between democratization and mass rule is most evident— and most fully explored by Goldfarb—in the institutions and practices of popular culture. Relying on work by the historian Donald Scott, Goldfarb suggests that the 19th-century transformation of the popular lecture from a community-based framework of civic education and cultural expression to a nationally organized format controlled and commodified by cultural entrepreneurs illustrates the development of popular culture in general. Mass culture is the soil that breeds cynicism; for mass culture is produced for "extra-cultural" (i.e., both pecuniary and ideological) purposes. Mass culture is, by definition, created cynically; it therefore socializes its consumers to have cynical expectations about public expression—and, finally, cynicism becomes the dominant tone and topic in mass cultural products. Yet, within every institution of popular culture, there remain democratic tendencies and practices—voices that are authentic, elites embodying values, publics that are autonomous.

Alongside mass culture, of course, has grown a mass politics that can be similarly analyzed. Political entrepreneurs on both the right and the left manufacture ideologies and techniques that attempt to construct mass constituencies. Right-wing ideology has been successful (and has deeply eroded authentic conservatism), the Left's move in this direction at the end of the 1960s (e.g., the Yippies) ended tragically. But, says Goldfarb (without much developed evidence or argument), a democratic public can be revitalized.

As is so often the case with books of insightful social criticism, this book is more likely to impress the reader with the diagnosis than with the proposed remedy. Goldfarb suggests that, to overcome cynicism, we (we intellectuals) ought to stop contributing to it. What he calls "ideological criticism," exemplified by Marxian and Mannheimian sociology of knowledge, is a major contributor to social cynicization. For ideological critique—the interpretation of cultural expression in terms of social interests—is no longer a game played by intellectuals; it has become a key resource for legitimating cynicism within mainstream cultural discourse. Instead of concentrating our energy on analyzing the social origins of ideas and of art, we ought to be evaluating their merits, helping ordinary citizens develop cultivated standards for assessing cultural quality and political worth.

I am troubled that Goldfarb wants to rely so much on a plea to intellectuals to believe in positive values and to abandon cultural analysis. For, as he at times seems to acknowledge, those who are socially subordinated really do need to know how their subordination is sustained by "ruling" ideas and symbols—and are, accordingly, aided by the kinds of criticism that disturb him. Indeed, a relentless questioning of the claims of those

in dominant positions—and of authority in general—seems essential for democracy. Only the systematic assumption that elites are not "sincere" gives the relatively powerless any chance to help shape the agendas of democratic debate. Goldfarb has helped us see that cynicism and reductionism have become resources for maintaining entrenched interests. This book, moreover, provides a model of how an academic sociologist can contribute effectively to democratic enlightenment. His leads, however, must be complemented by a continuing (albeit nonreductionist) analysis of the interplay between interests and ideas—between power and knowledge—if democratic culture is to be revitalized.

Critical Theory in Political Practice. By Stephen Leonard. Princeton, N.J.: Princeton University Press, 1990. Pp. xxix + 295. $35.00.

Anna Yeatman
University of Waikato

Critical Theory in Political Practice is a valuable intervention in what might be termed postmodern critical theory. Stephen Leonard attempts to recast the foundations of critical theory by detaching them from a modern(ist) "foundationalist" commitment to grounds of truth that lie outside the historical contingencies of human praxis. Leonard argues that Marx belied his epistemological insistence on the historicity of action and consciousness by his commitment to a universal philosophy of history wherein the proletariat is placed as the bearer of universal human emancipation. Equally the Frankfort school of critical theorists (Adorno, Horkheimer, Marcuse) maintained an orientation to a "normative universalism" (p. 257) by using this as a standard by which to assess how working-class and other movements have developed in this century. Habermas also maintains this commitment to normative universalism in the kind of consensual rationalism he views as the outcome of communication that is unconstrained by relations of domination. Leonard argues that Foucaultian postmodern critique negatively maintains this commitment in the universalism of its normative skepticism.

Leonard's argument is that the modern(ist) critical theorists, so skilled in critically discerning the historically contingent features of liberal universalist discourse, judged such contingency in the light of their own normative universalism. This contradiction has ensured that their emancipatory politics and critique paradoxically participate in an ahistorical and contemplative reading of human nature (natural law), albeit one quite different from that offered by the liberals and conservatives. It also ensures that their relationship to the historical contingencies of emancipatory politics and social movements is theoretical, always judging this politics in the light of some Platonic ideal. Once this ideal became irreparably tarnished by the complex realpolitik of 20th-century Communist states and Communist movements, the ideal counseled retreat into ideal-

istic despair rather than an active engagement with the contingent contradictions of contemporary praxis.

Leonard holds fast to the emancipatory theoretical tradition that the modern(ist) critical theorists developed. That is, he retains the core value that the role of critical theory is to assist human beings in their struggles against oppression and for self-determination and freedom from unnecessary domination. This depends on a rationalist assumption that a critically oriented rational, reflective orientation to practice provides the learning that must inform these struggles if they are to be successful. In turn, this learning is tested within a dialogical and participative relationship to critical theory's audience or constituency—whichever particular social movement context is at issue.

This is a rationalism that understands itself to be always specifically contextual and, thus, historically contingent. It has no difficulty, therefore, in reconciling the empirical existence of irresolvably different emancipatory struggles with the project of critical theory. This makes the project multiple and multivocal in character and constructs the histories, of which these projects are part, as open-ended as well as connected in a number of different ways. These connections arise, Leonard argues, because contemporary emancipatory movements that accept such plurality are also oriented to a politics of solidarity with each other.

Leonard particularizes and exemplifies this reconstruction of critical theory with reference to what he identifies as four "critical theories in practice": Cardoso's dependency theory, Freire's critical pedagogy, Latin American liberation theology, and feminist theory. He sees each of these as understanding themselves as reflectively related to historically specific contexts of practice. His readings of liberation theology and feminist theory are particularly compelling in these terms.

The task Leonard sets himself in this book is important, and he approaches it with courage and clarity. He shows himself to be aware of how this attempt to rework a rationalist emancipatory tradition in postmodern terrain may be criticized. For example the most likely criticism is that he jettisons the universalist presumptions on which his own attribution to the oppressed of an emancipatory politics depends. It is not clear that he has thought his way through this one. There are indications of a Christian emancipatory thematics that depends on an ontology of becoming and that invokes metaphors of redemption and witness, as well as speaking comfortably of "the" victims of oppression—or, "the" oppressed—who are to be assisted by critical theorists. These indications are particularly evident in the chapter on Freire; curiously enough, they are subject to more head-on critical awareness in the chapter on liberation theology, although it is not clear that the pastoral role of the critical theorist/theologian is critically confronted. Feminist theory actually occludes this problem by a complacent assumption that feminist theorists and their constituency are one and the same.

It is reasonable to accept the consequences of a nonfoundationalist rationalism for emancipatory politics, namely that it has no grounding

deeper than the historically expressed ideals and traditions of emancipatory movements. To do this I believe requires a critical theorist who takes the Nietzschean tradition of critique that informs Foucault's work more seriously than Leonard does. If critical theory is to become critical theor*ies* oriented within the secular and historically contingent pragmatics of particular emancipatory movements, then the question of the relationship of movement theorists to movement practitioners (let alone the wider, nonactivist movement constituency) becomes undodgable. Leonard argues that, if theory is to remain critical, it must maintain some distance in respect of the imperatives of strategic practice. This is true, but he does not inquire into what practical or institutional conditions make this possible, nor does he ask how it is that the distinctive will to power of critical theorists is to be made accountable to their practitioner constituencies. The point of Foucaultian critique is to show that all discourse—emancipatory or otherwise—is a discursive regime of inclusions and exclusions. How discursive experts can develop a critically reflective relationship to a regime that accords them discursive privilege is the kind of question that Foucault shares with late 20th-century critics of professional domination. It is a question that Leonard avoids. This said, this is a book that deserves more extended review discussion. It is a useful and serious contribution to those of us who are working with the conundrums of resituating modern(ist) critical theory in postmodern terrain.

Capitalism and Modernity: An Excursus on Marx and Weber. By Derek Sayer. New York: Routledge, 1991. Pp. x + 172 . $15.95.

Martin Riesebrodt
University of Chicago

Much of the debate in modern social sciences is a sometimes open, sometimes hidden argument about the relationship between the theories of Karl Marx and Max Weber. Many attempts have been made by Marxists as well as by Weberians to prove that these two classic theorists belong to different, if not opposite, camps. But there has existed from the beginning another tendency to stress the similarities rather than the differences in their descriptions of modern capitalism and identifications of the problems of modernity. Derek Sayer's book belongs to this tradition, one with which I agree as long as the differences between Marx and Weber are not overlooked.

Capitalism and Modernity, therefore, is a very welcome attempt to remind us of the parallelism in the analyses and criticisms of modern capitalism by Marx and Weber. The book is divided into four chapters. The first two deal with Marx's treatment of capitalism and the state, the second two, with Weber's concepts of modern rational capitalism and bureaucracy. Sayer certainly succeeds in demonstrating striking parallels

in the writings of both authors. But the overall picture of the book suffers somewhat from the contrast between the high expectations raised by an exaggerated cover text and the actual book, which, from my perspective, has some shortcomings, especially when compared with Karl Löwith's classical study *Max Weber and Karl Marx*.

Sayer is a Marx specialist. But despite his impressive familiarity with the texts, he does not address the problem of how to interpret them. On the one hand, Sayer's unqualified use of Marx's writings from all periods of his work suggests that his reading is based on the assumption of continuity between the early and the later works. On the other hand, his neglect of the concept of "alienation" seems to indicate the opposite. He fails to address the problem even implicitly and causes confusion about his position.

In refuting criticism of Marx, Sayer's strategy seems not completely adequate. It is insufficient to prove that Marx did not believe in techn(o-log)ical determinism, or did not underestimate the (relatively autonomous?) role of the modern state by a simple quotation. There are also passages in Marx's writings that lend support to those interpretations, and Sayer, in my view, should have made a better argument to defend his reading of the texts. This is unfortunate, because Sayer's mastery of Marx's writings is impressive and competent.

This, however, is much less true with Weber. Sayer's genuine interest in Weber seems to be rather limited. He uses him primarily to legitimize and, where necessary, supplement Marx. This limited interest leads to an obvious imbalance between the different chapters and to a rather reductionist reading of Weber, stressing primarily his economic interpretations of history and modernity. Given the rich recent literature of Weber's understanding of Western development as expressed in concepts such as institutional and ethical rationalization, calling, and personality, Sayer's interpretation of Weber is less sophisticated than much of the present debate. Although some of this literature is included in Sayer's bibliography, it seems to have had little effect on his arguments.

The book claims to contribute to a better understanding of our modern condition in the late 20th century. But this claim is hardly validated. Sayer's conclusions are moralistic rather than analytical, and surprisingly indeterminate. Sayer wants to be at the same time a feminist, anti-"Orientalist" critic of Marx and Weber as well as a Marxian and Weberian critic of capitalism and modernity. Sayer agrees with criticisms of Marx and Weber as writing from a "male" and "Occidentalist" perspective. Even if these criticisms were correct, what would their implications and consequences be? Does this disqualify the analysis of Marx and Weber or not? If not, why not? If yes, in what respect? If the claim of the uniqueness of Western development is based on an "Occidentalist fallacy" then we have to give up the diagnoses of Marx and Weber. If this claim, however, is a useful and justifiable one, then we have to give up or at least modify the criticism. One cannot have it both ways. The same holds true for feminist criticism.
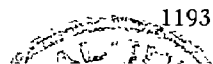
Sayer tends to accept all kinds of fashionable criticisms uncritically and without drawing any theoretical consequences. And he does not reflect on the changes within modern capitalism since Marx and Weber. Why does he believe that their analyses and criticisms still apply despite their bias and the structural changes? Sayer does not address these and other issues. As a consequence, he does not succeed in linking his discussion of capitalism and modernity with a critical diagnosis of our times. Although Sayer has written a partially stimulating and interesting introduction to an important aspect of classical social theory, Karl Löwith's study still remains unsurpassed as a sophisticated comparison between Marx and Weber that treats both authors in their own right.

Structures of Social Life: The Four Elementary Forms of Human Relations. By Alan Page Fiske. New York: Free Press, 1991. Pp. xii + 480. $29.95.

John F. Stolte
Northern Illinois University

Suppose a group, for example a small town, faces a problem, say, providing for fire fighting. An observer might divide this problem into different social relational components: social decision making, moral thinking and argument, the social organization of work, the production and distribution of resources, or the motivations, meanings, and judgments of the group's members. Alan Fiske argues that each such component might be structured in terms of any one or a combination of four basic models: communal sharing (CS), authority ranking (AR), equality matching (EM), or market pricing (MP).

For illustration, imagine how fire fighting might be socially organized as work. Modeled on communal sharing, fire fighting would be seen as a collective responsibility, with each townsperson contributing what he or she could to the service and using what was needed of the service. Authority ranking would imply superordinate townsfolk controlling the fire-fighting activities of subordinate townsfolk. In equality matching, each townsperson would perform exactly the same fire-fighting activities, perhaps by taking turns. Under market pricing, each fire fighter would perform the service in exchange for a formally calculated wage determined in a labor market.

In a broad synthesis of theory and research on many topics, from resource exchange through social choice and social motivation to ideology and legitimation, the author discerns the appearance and reappearance of the same four underlying models, sometimes partially stated, sometimes conflated. Across the writings of Weber, Piaget, and Ricoeur (as well as many others), Fiske sees underlying parallels, similar themes independently invented and reinvented. These conceptual convergences together with ethnographic observations Fiske made of a West African people,

the Moose of Burkina Faso, suggest to the author that there are just these four basic models, which are *fundamental* (each is a basic grammar for social interaction), *general* (they cover most social interaction, thinking, and feeling), *elementary* (they are simple building blocks from which complex social structures are composed), and *universal* (they operate in all persons in all cultures). Indeed, in Fiske's view, the models are so basic that they must be psychologically endogenous, something like Chomsky's biologically given grammatical knowledge that "humans universally share at birth" (p. 187).

The models are invariant, it is argued, but become specified in varying ways across cultures because of "implementation rules." Such rules stipulate the social domain, relevant actors, parameter settings, relational marking code, and ideological judgments that qualify, restrict, and shape concrete manifestations of the four models.

Social order emerges, says Fiske, at different levels of complexity (dyad, role sets, organizations, institutions, societies) through the recursive application of one or more of the models in nested hierarchies at different levels of inclusion.

By way of general appraisal, one sees impressive strengths and serious shortcomings in this work. On the positive side, most sociologists will find Fiske's analysis of how the models shape the justification of misfortune and the perception of time very innovative and interesting. Also, his proposed ontogeny of the models (an invariant order of stages as follows: CS, AR, EM, and MP) and the mathematical/measurement scale properties of the models (CS = nominal, AR = ordinal, EM = interval, and MP = ratio, etc.) will be judged brilliant. Finally, his adroit demonstration that Western social science has been far too preoccupied with the "rational choice" model, which is, in fact, quite narrow and ethnocentric, will be loudly applauded by many.

On the negative side, Fiske's argument will be roundly criticized for its overly functionalist cast, despite his disclaimers to the contrary. One indeed wishes the social world had as little conflict, inequality, and "raw power" use as the author seems to think it has. Most sociologists will simply not settle for relegating such pervasive realities into a "null" category of "asocial" relations or "sociopathy." Further, Fiske's attack on the traditional notions of "norm internalization" to buttress his own concept of "model externalization" is simply exaggerated. Surely there is a compromise between these principles that is closer to the truth. Finally, and most damning, Fiske's argument, at bottom, is a radical cognitive reductionism, despite all the arm waving about "emergence" and "social networks." In the 1960s and 1970s, Homans and the radical behaviorists failed in their attempt to reduce complex social interaction events to individual, operant conditioning. Today, Fiske fails to reduce such events to a priori, individual cognitive models. In neither case are the links and processes that connect individuals into simple or complex social structures adequately laid out.

In sum, despite several important psychologistic shortcomings, this

provocative, beautifully written book is worth serious sociological attention.

In Search of Human Nature: The Decline and Revival of Darwinism in American Social Thought. By Carl N. Degler. New York: Oxford University Press, 1991. Pp. xi + 400. $24.95.

Stephen K. Sanderson
Indiana University of Pennsylvania

Carl N. Degler's *In Search of Human Nature* tells an important story in the history of social science: social scientists' retreat from, and eventual return to, the role of human biology in the explanation of human social behavior. According to Degler, in the late 19th and early 20th centuries, social scientists widely agreed that much of human behavior was rooted in biology. The biological factors of race and sex were considered major determinants of behavior. Racism and sexism were pervasive: nonwhites, immigrants to America from southern and eastern Europe, and women were regarded as inferior to white males of Anglo-Saxon descent. The widespread belief in genetic differences among social groups spawned the eugenics movement, which was led by Francis Galton, Charles Darwin's cousin.

An antibiological reaction then set in, led by such scholars as the anthropologist Franz Boas. Boas radically separated the concepts of race and culture and argued that culture was an entirely extrabiological phenomenon. Boas greatly influenced such social scientists as Alfred Kroeber, Robert Lowie, and Howard Odum, and by the 1920s these scholars and others succeeded in winning priority for their argument that differences between races were to be explained entirely by environmental influences. An intellectual shift with regard to sex differences followed much the same path as that for racial differences. The argument that sex differences were fundamentally environmental rather than genetic eventually won the day by the 1930s. The antibiological forces also fought and won battles with respect to several other aspects of human behavior: earlier attempts to link crime to the allegedly inherited trait of feeblemindedness were successfully attacked; the concept of instinct widely employed by psychologists was driven from the scene by behaviorists like John B. Watson; and the eugenics movement was severely criticized and had declined greatly in popularity by 1930.

Why, by the 1930s, had the antibiological forces triumphed, and why was their victory achieved with such surprising ease and in such a short time? Degler's answer is that new empirical evidence had little to do with it. The data were sufficiently limited and ambiguous so as to preclude a clear victory on strictly scientific grounds by either side. The triumph of the antibiological forces was largely ideological, Degler claims. The psychologists, anthropologists, and sociologists of these days had strongly

egalitarian values and reformist goals, which they perceived to be severely threatened by an emphasis on biology as a determinant of behavior. Environmental explanations of behavior, on the other hand, fit beautifully with social scientists' ideological outlook, and, in the absence of empirical data pointing strongly one way or the other, such explanations were embraced.

By the 1950s and 1960s, however, some social scientists began to feel that social science's antibiological stance had gone much too far. As a result of dissatisfaction with prevailing theories and the emergence of new developments in the biological sciences, some social scientists wanted to reconsider biology's role in human action. Psychologists led the way, with sociologists and anthropologists getting into the act somewhat later. The ethological work of thinkers like Nikolaus Tinbergen and Konrad Lorenz was applied to humans by, inter alia, Desmond Morris, Lionel Tiger, and Pierre van den Berghe. The sociobiology of Edward O. Wilson became prominent and influential after the mid-1970s. Degler traces in detail, and with considerable approval, the application of sociobiological theories to such dimensions of human social behavior as incest avoidance, social hierarchies, altruism and reciprocity, and male-female differences in aggression and parental care. He insists, correctly in my opinion, that the revival of biology in the social sciences has nothing to do with the embracement of racism, sexism, and other social evils, as is feared by biology's opponents.

Degler's excellent book deserves a wide reading by social scientists because it tells an extremely important story about sociological influences on social science. Its most important message is that social scientists' antagonism to considering the biological foundations of human behavior is rooted far more in ideology than in science. Environmentalists won the debate with hereditarians in the 1920s and 1930s because of reformist zeal rather than empirical evidence, and the current antagonism to sociobiology on the part of so many social scientists stems largely from both their political ideology (strong commitment to egalitarianism, fear of racist and sexist misuse of sociobiology) and their disciplinary ideology (widespread adherence to the Durkheimian dictum that "social facts can only be explained in terms of other social facts"). I predict, though, that the time will eventually come when antibiological social scientists can no longer explain away the mass of data demonstrating important biological foundations of human behavior that will have accumulated. When that day comes, many social scientists are going to have egg on their faces.

Ancient Cultures of Conceit: British University Fiction in the Post-war Years. By Ian Carter. New York: Routledge, 1990. Pp. x + 317. $54.00.

Elizabeth Long
Rice University

Rarely can a sociologist be accused of penning a witty romp. *Ancient Cultures of Deceit,* however, is just that—an enormously readable book by sociologist Ian Carter about British university fiction and the fate of British higher education in the past 20 years. It is also an engaging example of the little-known genre "sociologists fight back." Carter was first intrigued by these novels because of their vilification of sociology. His revenge on the fictioneers is to hoist them by their own petards, subjecting them to a deft dissection, while beating them at their own game of ironic intertextuality.

Defining British university fiction as a "cross-generic genre" recognizable when a novel's action is "centred in a university" (p. 4), Carter analyzes the discourse this fiction expresses. According to him, it has two leading features: "First, Oxford and Cambridge, the ancient English universities, form the template against which other kinds of universities are measured and found wanting. Second, the university is treated as culture's citadel, besieged by four barbarian hordes: proletarians, women, scientists, and foreigners" (p. 215).

The chapters detailing these claims are among the strongest and most pleasurable in Carter's book. He demonstrates that every level of novelistic representation—from passages extolling the bells and towers of Oxbridge colleges and their gently cultured dons to those bemoaning the earnest frumpiness of Oxbridge females or the totally hopeless vulgarity of non-Oxbridge universities—works to induct the reader into a very particular understanding of civilized "culture." In this discourse, culture signifies aristocratic Englishness rooted in social and masculine privilege, and learning signifies genteel humanism. Novelists portray elite universities as "finishing schools for men born to rule" (p. 83), while summarily dismissing other educational institutions, other social constituencies, and other intellectual priorities.

More problematic are Carter's strategies for demonstrating that this discourse has affected the politics of higher education in Britain. One problem is that neither his theoretical framework nor its empirical dimensions are systematically elucidated; indeed, his ironic style would make this difficult. His basic contention is that the discourse of British university fiction helped subvert the very values it meant to protect: "tolerance, decency, moderation, democracy" (p. 263). Its misplaced defense of higher education against an imagined threat from the left and the lower orders enabled Thatcher's minions, the true "Ostrogoths," to dismantle the entire system according to the very "technologico-Benthamism" dreaded by conservative humanists, including authors of British university fiction.

This seems a plausible argument, but the interconnections between fictional discourse and political action remain tenuous. For instance, Carter outlines the conservative attack—in the infamous *Black Papers* and other publications of the 1970s and 1980s—that first challenged the expansive postwar program of British higher education. Connecting its leaders and their ideas with F. R. Leavis and his journal *Scrutiny,* Carter links this movement to British university fiction by characterizing both as rightist interpretations of Matthew Arnold's ideas about culture and education. Again, this is plausible, but in the absence of empirical evidence about the processes of political decision making in Britain, little has been shown about the political effects of fictional discourse.

Part of the problem arises from Carter's failure to exploit fully the theoretical or empirical approaches adumbrated by the term "discourse," so his cultural analysis resembles the "images of . . . in fiction" approach with its attendant problems of connecting literature to the social world. The more social-constructionist framework implied by discourse reaches beyond literary texts by several routes. Crucially absent here are a concentration on audience response to the novels or their television adaptations mentioned by Carter (an important step in investigating a discourse's effects), and a fully rendered discussion of the environing discursive field, usually construed in this perspective as a contested domain of social/symbolic practices engaged in by groups with different social positions and interests.

Carter's book does not make it clear to me, an American reader, who the relevant social actors were and whether or how public "consent" (and if so, whose) to Thatcher's program was mobilized, let alone specifically how British university fiction fed into this process. Equally important, Carter shortchanges discursive influences other than that of Matthew Arnold. Yet his own narrative makes it clear that the Thatcherites eschewed Arnoldian humanism. Similarly, Raymond Williams and the beleaguered London universities and departments Carter discusses were close to Marxism and other strands of critical theory, yet this is never mentioned. And, finally, perhaps because Carter does not fully explore contestation either within or beyond the universe of British university fiction, his sense of alternatives to Thatcher's programs is underdeveloped. He proposes a return to the Scottish system, grounded in moral philosophy and the social sciences and more socially open than the English system, instead of also summoning possibilities from within existing British institutions or from beyond the British Isles.

Nonetheless, Carter's account of the right-wing evisceration of the British university system is instructive and eerily timely. And despite the issues raised above, his take on British university fiction is a delight: a provocative critique of ideology that, by implication, invites a similar examination of American intellectuals' self-serving notions.

Neighborhood Tokyo. By Theodore C. Bestor. Stanford, Calif.: Stanford University Press, 1989. Pp. x + 347. $39.50 (cloth); $11.95 (paper).

Karen E. Campbell
Vanderbilt University

In *Neighborhood Tokyo,* anthropologist Theodore Bestor explores a small residential and shopping area bordering on central Tokyo; he calls it "Miyamoto-chō" and assures us that it is typical of other small Tokyo neighborhoods. In addition to providing a detailed ethnography of this unremarkable neighborhood, Bestor identifies two analytical goals. He aims (1) to debunk common views of Japanese neighborhoods— particularly the notion that patterns of life in contemporary neighborhoods are simply continuations of village traditions—and (2) to consider the importance, nonetheless, of traditionalism in shaping neighborhoods over time (pp. 8–11).

In chapters 3–7, Bestor explores several dimensions of Miyamoto-chō: how the neighborhood fits into larger political structures, informal services provided by the *chō-kai* (neighborhood association), the formal status hierarchies linked to the *chō-kai,* informal networks in Miyamoto-chō, and the role of festivals in local life. Chapter 6 ("Friends and Neighbors"), in particular, should be widely read by those curious about informal networks and their functions in cultures beyond those of the United States and Canada.

As an ethnography, the book is quite effective: it is well-written and pleasant to read, filled with a myriad of details about life in Miyamoto-chō. What puzzled me was the virtual absence of the author himself (who, with his wife, lived in the neighborhood for several years). But for a handful of illuminating anecdotes, the story of Miyamoto-chō might have been recounted by a wraith who, unnoticed, listened in on conversations, meetings of the *chō-kai,* and festival planning. I would have been reassured had Bestor shown more cognizance of the possible effect of his presence on everyday interactions. Nonetheless, he succeeds in his goal of providing "a detailed ethnography of . . . the texture and structure of social relationships" in Miyamoto-chō (p. 9).

Bestor is less than convincing in his argument that Japanese neighborhood life is shaped largely by the activities and interests of the "old middle class," an assertion he makes early in *Neighborhood Tokyo,* but does not explicitly return to until the concluding chapter. True, the prime actors in Bestor's descriptions of political activities, festival planning, and daily life are local business owners and their families. He argues (p. 264) that these residents effectively define the neighborhood and status within it in ways that exclude white-collar workers employed beyond the neighborhood. But the book lacks clear evidence that such *sarariimen* are, in fact, so excluded—their simple absence from Bestor's descriptions is not in itself sufficient.

The most intriguing and intricate analysis in *Neighborhood Tokyo* re-

volves around the disputed role of tradition in contemporary Japanese communities. In chapter 2, Bestor traces in considerable detail the history of the neighborhood, laying the foundation for his assertion that current community patterns cannot be interpreted as extensions of centuries old rural traditions. The area that now comprises Miyamoto-chō *did* begin life as a rural village, but—if Bestor's history is complete—its saga is characterized more by upheaval than continuity. First, the community was transformed during the 1920s and 1930s by large numbers of out-migrants from neighboring Tokyo; the population of the area surrounding Miyamoto-chō grew more than 15-fold between 1920 and 1930. A second transformation occurred around World War II. During the war, over half the area's residents were killed and 40% of the housing stock was destroyed. These cycles of rapid population expansion and contraction and waves of migrants from the urban center, Bestor concludes, mean that today's Miyamoto-chō is a far cry from the rural village that previously occupied this plot of land.

Thus, Bestor early in the book casts doubt on the vision of Japanese neighborhoods as extensions of feudal beginnings. Throughout *Neighborhood Tokyo,* it is clear that tradition is an important facet of life in Miyamoto-chō, evidenced in the activities of the volunteer fire brigade, an annual festival centered about the local Shinto shrine, and ritual gifts to families and institutions. But, Bestor argues, tradition is only rarely a direct link to the rural past. Rather, the "old middle class" uses tradition to sanction current neighborhood activities, and the community as a whole apparently accepts this mutable vision of tradition. The traditionalism of those who dominate local politics and social services is "an active agent creating and re-creating images and meanings that attach to the social forms around which people organize their lives" (p. 262). in the end, then, Bestor turns the usual interpretation of tradition on its head: instead of limiting local activities and relations to forms carried through the centuries, tradition (in the hands of community leaders) serves to legitimize changes that facilitate neighborhood survival in an urban setting.

Sporting with the Gods: The Rhetoric of Play and Game in American Culture. By Michael Oriard. New York: Cambridge University Press, 1991. Pp. xviii + 579. $49.50.

Norman K. Denzin
University of Illinois at Urbana-Champaign

This is a massive work, its sweep is majestic. It attempts nothing less than a reading of the past two centuries of American life through the metaphor of life as a game. Three key terms (play, game, and sport) and their transformations (amusement, mirth, gaming, gambling, and sportsmanship) define the inquiry. The analysis moves from 19th-cen-

tury play, sport, and western mythmaking (those cowboys in Owen Wister's *The Virginian*), to southern honor (that famous hospitality, fox hunting, and holiday frolics) and gender (those duels), to sportsmanship and gamesmanship in 20th-century fiction (Faulkner, Fitzgerald, Hemingway), to play and counterculture in the 1920s, ending with the beats, the hippies, and the New Age followers of the 1980s.

The thesis is straightforward, yet complex: "Play is a projection of a society's inner life" (p. xi). Accordingly, key American values are given in the shifting political, ideological, and cultural meanings brought to the rhetorics and discourses surrounding the play and gaming experience. These values and their transformations follow, in a complex fashion, the three stages of capitalism in American life (market, monopoly, and consumption). A progressive devaluation of work accompanied these shifts, giving rise to a profound separation of capital from labor, with new forms of play and gaming fitted to the leisure and working classes. This separation of capital from labor paralleled the play-work ideologies of the 19th-century transcendental movement and its theology of sentimentalism (Emerson, Thoreau). The Progressive Era, which extended from the 1880s to World War I, elaborated these themes, making the 19th century America's century of work, and the 20th, its century of play (p. xi).

Yet Americans have always had profound doubts about pure play (the legacies of the Protestant ethic), preferring to mix play with work. The game metaphor permitted this mixture. If play is activity for its own sake, and work is goal oriented, then a game embodies both activities and both motives; hence the game of business, the games of professional sports, and so on. In the gaming metaphor, which came to stand for the game of life, Americans had their play, their business, their profit, and their work, all in the same experience. More important, gaming came to represent an arena where central cultural values—heroism, honor, valor, dignity, strength, honesty, and hard work—could be enacted and vicariously experienced. Encased by political rhetoric, and capable of being fitted to any slice of life, these three terms (play, work, and gaming) became metaphors for life itself. Warfare and capitalist business ventures are games conducted under the rules of fair play, and ordinary citizens are encouraged to play at their work. Hence the title of Michael Oriard's book for, in playing and gaming, Americans have dramatically enacted their particular and multiple versions of sporting with the gods.

The textual materials for this cultural study lie in the author's reading of mainstream, "classic" American literature, including the writings of Emerson, Thoreau, Hawthorne, Melville, Twain, James, Hemingway, Faulkner, Fitzgerald, Ginsberg, Kerouac, Bellow, Mailer, Salinger, Pynchon, Kesey, and Updike. Mindful of the limitations and biases in these works, Oriard reads them conflictually, creating spaces for the many marginal fictional voices, including the feminist (Jong, Brown, French, Daly) and the African-American (Baldwin, Ellison), that challenge these canonical texts. In these literatures he discovers a society at odds with

itself. The "linguistic constructions of life as game or play" (p. xv) simultaneously reinforce, resist, enhance, contradict, and, sometimes, consolidate existing power and ideological arrangements in the shifting main and marginal streams of American culture.

Since the mid-1800s the problem has remained the same. Americans have been unable to satisfactorily "ground American social reality in a . . . satisfying integration of work and play" (p. 484). From the bohemians of the 1920s to the beats of the 1950s, the hippies of the 1960s and the New Age romantics of the postmodern 1980s, new, utopian sensibilities stressing the discourses of play and freedom interact and conflict with dominant cultural values stressing religious piety, hard work, sacrifice, and conspicuous consumption.

The visionary impulses of feminists, critical theorists, and Marxists persist. Utopians seek "a new self that coexists harmoniously with all other life" (p. 484). They await a totalizing social movement that will finally break down the competitive, exploitative boundaries between the self and the world. But the dream of such a movement has been constant throughout America's history. The current postmodern, yuppie, New Age phase is only the most recent marker in a long line of previously failed attempts to meaningfully construct an image of life that, guiltlessly and effortlessly, weaves the play elements of the self into a peaceful, unified picture of human group life.

All is not lost. Oriard wants us to believe that Americans and their writers have become more sophisticated about the rhetorics that surround play and work in their culture. He argues that "people can only realize what they first imagine . . . there is hope in the fact that writers have become more . . . self-conscious in addressing" these persistent problems (p. 484). But what is imagined is itself determined by material conditions, power arrangements, and the communications industries that bring these issues in front of us. Until there are fundamental and drastic changes in the cultural logics of late capitalism, American society will continue to get the kind of play and games it deserves. And this is tragic, for the God called capitalism that Americans do their sporting with is not joyful, free, gentle, kind, or playful.

A minor point. A missing text in Oriard's book is Leslie Fiedler's *Love and Death in the American Novel* (New York: Stein & Day, 1986). Regarded by critics as one of the best readings of American literature in the 20th century, Fiedler's book covers the same time period and the same literary texts as *Sporting with the Gods*. But Fiedler discovers a pattern in American culture that is missing in Oriard's analysis; namely the inability of the white male American fiction writer to deal with "adult heterosexual love and his consequent obsession with death, incest and innocent homosexuality" (Fiedler 1986, p. 13). Fiedler finds this theme embedded in virtually all American literature, from James to Faulkner, Hemingway, Fitzgerald, and Chandler.

Fiedler's argument is by now familiar. But his readings suggest that another metaphor, darker, and more fatalistic, lurks behind American's

play and games. Fearful of the other called "woman," American males (and their writers) seek a fiction that romanticizes love, yet is unable to deal with concrete, loving, intimate relations with women. Hence the attraction of sports and male games and male bonding. In the pseudo-homoerotic, sporting buddy relationship, men can have their honor, love, and intimacies without the risk of a commitment to an other who remains elusive, frightening, and terrifying in her emotional demands and interactional possibilities.

The rhetoric of play and game in American culture is about sporting activities and games with two gods (or goddesses): capitalism and women. Oriard has performed a brilliant analysis of the first game, but has failed to unravel this other goddess who is hidden inside the masculine structures of the metaphor that says life is a game.

Backboards and Blackboards: College Athletics and Role Engulfment. By Patricia A. Adler and Peter Adler. New York: Columbia University Press, 1991. Pp. xii + 262. $35.00 (cloth); $17.50 (paper).

Thomas J. Cottle
Massachusetts School of Professional Psychology

That college basketball is a billion dollar business is known to every American sports fan. That certain institutions of higher learning have become the breeding ground, or minor league, for professional sports is also well known. Not that well-known, however, is the perplexing and often sad manner in which American universities deal with the young men who presumably have entered their gates to become educated, but who clearly matriculate only to play ball, which in turn fills the university's coffers and perpetuates the dreams of ambitious men and college boosters who, seemingly, live and die with the successes and failures of very tall teenagers.

In *Backboards and Blackboards,* Patricia A. Adler and Peter Adler have written a participant-observer study of a high-powered division 1 university basketball team, tracing the lives of coaches, players, and boosters. Basing their research on fundamental sociological concepts like Glaser and Strauss's grounded theory and Blumer's notions of adaptation, and carefully defining the various role activities of athletes, the Adlers portray one of the ignoble chapters of contemporary American life. Even more, they provide insight into why it is that so many people are seduced by the world of national sports and the sheer magic of hanging out with kids who might someday be touched by the heat of professional stardom.

The book follows a southwestern university's basketball team, examining the recruitment of players, the hierarchy of team members ranging from the headier to the more physical and tough ones; the role of the coach, a man who serves as father, teacher, academic broker, and, at

times, (ill) adviser; and finally the everyday lives of young men, many of them totally unprepared for the social, intellectual, and even athletic responsibilities to which they have committed themselves.

Leaning heavily on the concept of the student's role, be it social or academic, glorified or mundane, Adler and Adler codify and systematize daily activities as well as the expectations placed on these few extraordinary students, very few of whom ever adapt or are helped to adapt to the academic demands of the university. Mostly as sociologists, but, occasionally, almost as sympathetic parents, they record the lives and comments of young men glorified by the local media and fans. All of them are engulfed in an undoubtedly exciting but often exploitative landscape meant ultimately to earn money for the university and its community, bring glory to its students and alumni, and keep alive for a few short years the mainly unrealistic dreams of mostly poor young men.

It is around the concept of role engulfment that much of the research centers. The essential identity of these young players is carved out of the expectations others have of and for them, but in this unique case, these expectations and attendant roles bedazzle everyone. Simultaneously, they learn that they cannot (or need not) commit themselves to classroom work and that everyone identifies them as basketball players, no matter what they do and no matter what they think of themselves. Indeed, as the authors eloquently report, the players develop a master's status within the community, a poignant status, actually, given that within a few years they will be obliged to disengage from the university and terminate all facets of their evolving self concept—if, in fact, they actually graduate— along with practically every activity associated with the role they now play (pun intended).

Backboards and Blackboards is good, strong, in-your-face sociology, but its greatest contribution well may be captured in an opening (and previously published) chapter in which the junior author offers the reader a taste of what he experienced as team adviser. In these pages, one finds the seductive aura of college athletics, the lure of locker rooms, the thrill of rubbing elbows, as they say, with young athletes whose celebrity radiates a magical electricity. The hopeful dreams of boyhood and the unfilled ones of manhood begin to be heard in this chapter, and one recognizes just why it is that presumably honorable institutions openly corrupt the values and lives of students, and, in the process, ironically (and cynically) train them for the self-glorification, personal merchandising, public-relations superficiality, and the devalued currency of intellectualism that these young people will find in the so-called real world.

According to the authors, the world of college athletics is far more backboards than blackboards. I found the volume utterly competent in traditional sociological terms, but its subtext painful and pathetic for what it says about American culture and its insatiable need for celebrity, entertainment, and distraction. That colleges condone this sort of exploitation and look away from the unethical behavior of faculty—when the faculty are coaches—is pitiful but, alas, predictable. At least now it is

documented and framed in orthodox and readable sociological contexts. Moreover, it is given life through the words of young men whose court feats, presumably, will not be forgotten on this one particular campus, but whose lives were hardly enhanced by a college population that—with the one exception of their girlfriends—only pretended to value them.

Getting Started: Transition to Adulthood in Great Britain. By Alan C. Kerckhoff. Boulder, Colo.: Westview Press, 1990. Pp. xviii + 216. $35.00.

Adam Gamoran
University of Wisconsin—Madison

Getting Started concerns the interaction of educational, family, and work careers, during late adolescence and early adulthood, for a 1958 British birth cohort. It purpose is "to describe the British pathways, to show how they differed from those in the United States, and to illuminate the role of the two educational systems in bringing about those differences" (p. 177). The book adopts a life-course perspective in its attention to the timing and sequence of school leaving, labor-force entry, moving away from home, marriage, and parenthood. In addition, it shares the status-attainment perspective's concern for the distribution of these events across the British social strata.

Alan C. Kerckhoff has written extensively about stratification in Britain and the United States, and the educational system is prominently featured in much of his work. A major contribution of the present volume is an examination of the distinct effects of regular secondary and higher education, on the one hand, and largely part-time "further education," which can provide additional educational certification, on the other. The distinction was motivated by an effort to understand an unexpected conclusion of earlier studies: despite a "sponsorship" system, in which British youth were divided into selective and nonselective secondary schools, overall the intergenerational transmission of status was no higher in Britain than in the United States. Perhaps, Kerckhoff reasons, widely available opportunities for further education were compensating for status reproduction that occurred through the sponsorship system. This issue and others are addressed with the National Child Development Survey (NCDS), a survey sample that began with virtually all children born in England, Scotland, and Wales during the week of March 3–9, 1958. Data on school and work experiences were subsequently collected at ages 7, 11, 16, 18, and 23.

Getting Started is written for an American audience, and its author makes no assumptions about the reader's knowledge of the British educational system. Rather, he provides a helpful description of the organization of schooling in Great Britain, as it existed for the subjects of the study (i.e., through 1981). There are several essential differences between the British and American school systems. (1) In addition to comprehen-

sive secondary schools, as in the United States, there were high-status grammar and private schools, with selective entry criteria, and lower ranked secondary modern schools. Moreover, ability grouping in British secondary schools was more formalized than American tracking systems. (2) Whereas educational attainment in the United States is appropriately measured by years of schooling and diplomas obtained, the British have an elaborate system of qualifications that one acquires mainly by passing examinations at particular time points. One's qualifications at the time of initial labor-force entry are heavily influenced by one's school type and ability group during the secondary years. (3) British students tended to complete their regular schooling at an earlier age than Americans (medians of about 16.4 years old in Britain and 18.4 in the United States). However, over a quarter of the British students obtained additional qualifications through further education, whereas Americans less often obtained additional certification once they left school.

Partly as a consequence of these schooling differences, the sequence of life-cycle events is more compact in the United States. School completion, labor-force entry, and leaving the parental home tended to occur within a six-month period or so, with marriage following at about age 20 for females and 22 for males, and parenthood typically arriving at around age 24.5 for females and later for males. The British tended to leave school and begin working at an earlier age and to leave home, marry, and have children a year or two later than their American counterparts.

The book's most important findings concern changes in educational qualifications and occupational prestige that occur between the time of first job and age 23. Is further education an alternative route to high-status jobs? Kerckhoff explains that if further education is dominated by those from advantaged families and selective schools (as is regular education beyond age 16), then it would reinforce the inequalities of the sponsorship system. If, however, further education is populated more by those from low-SES families and low-status schools, then inequality would decline as a result. When work experiences are controlled, Kerckhoff finds no SES differences in the likelihood of obtaining additional qualifications after labor-force entry. In addition, school type has no direct effect on additional qualifications. Ability grouping has no direct effect on extra qualifications for females, but among males, high- and middle-group students are advantaged. Overall, in contrast to regular schooling, further educational opportunities do not add to the advantages of those from wealthier families and selective schools and groups nor, however, do they compensate for inequalities that exist at the time of initial school leaving. Hence, although further education provides an alternative pathway to educational qualifications, it cannot account for why the sponsorship system of ability grouping and selective schools does not yield a higher degree of status inheritance.

For students of status attainment and of the late adolescent/early adult life cycle, this book provides valuable information. I recommend it especially to researchers of comparative educational stratification, for whom

it is essential because of the evidence on the contribution of further education. Finally, I recommend reading the book to become acquainted with the idea of a birth cohort as a longitudinal survey sample, for such a sampling frame offers advantages over the school-based national educational data sets we have in the United States.

Young Mothers? By Ann Phoenix. Cambridge, Mass.: Basil Blackwell, 1991. Pp. 267. $59.95 (cloth); $26.95 (paper).

Arline Geronimus
University of Michigan

Young Mothers asks, "Are teenage mothers too young to be good mothers?" while implicitly asking, "Is existing evidence sufficient to support the stereotypes and conclusions maintained about the 'problem' of teenage motherhood?" Ann Phoenix's answer to both questions is no. The empirical basis of her book is a longitudinal study of a group of English teenage mothers. Phoenix interviewed 50 sample members first in late pregnancy, then six months after childbirth, and again as their children turned two (more were interviewed fewer times). Developmental tests were administered to the children at age two. The in-depth interviews attempted to document the women's life histories and their experiences as mothers.

The book has introductory chapters and an endnote with chapters on how the women came to be mothers, men and marriage, social networks and emotional support, practical support, the children, and employment, education, and satisfaction with motherhood. Phoenix laces the book with interview quotes and summary impressions including a critical review of relevant literature and insights, anecdotes, and wry observations in support of her arguments. Overall, she finds the outcomes for young mothers and their children unremarkable; to the extent that some do have palpable, primarily financial, problems, she interprets these as attributable not to maternal age, but to factors that preceded pregnancy such as working-class backgrounds, lack of academic inclination, or employment in jobs devoid of training or career prospects. Some who advanced beyond these inauspicious beginnings reported doing so *because* they had become young mothers and were motivated by the desire to provide better for their children. Her informants saw motherhood as an important career, had long anticipated becoming mothers young and felt that motherhood would be the most fulfilling aspect of their lives. Despite suboptimal circumstances for administering the developmental measures and their possible class biases, the scores of most children were above the norm.

The empirical evidence raises intriguing questions on marriage. This is important because of the easy link between early and out-of-wedlock childbearing and the belief that marriage is an ideal source of social and

financial support for mothers and children. The informants told a less romantic story. More than a quarter were unequivocally negative about marriage, while almost 40% felt men benefited more from marriage than women. They perceived women's role in marriage to entail drudgery and were more likely to say that marriage would curtail their freedom more than having children. Married respondents reported their husbands to be uninvolved in child care or household chores. Phoenix notes that male partners who were married or cohabiting with the mothers were less likely to be unemployed than nonresident fathers, leading her to speculate that young mothers follow the normative prescription of marrying only when their mates can be financially supportive. In contrast, informants reported networks of female friends and relatives providing substantial amounts of social and practical support. Phoenix observes that the women's ability to maintain such networks was threatened by serious involvements with men or marriage. These observations raise questions about whether teen mothers trade marriage for other (better) sources of support and about how to view marriage and social or practical support for motherhood in working-class and poverty contexts. Systematic answers to these questions would have paradigmatic and policy implications.

While this book provides provocative and helpful descriptive information, ironically the empirical portion suffers from all of the major methodological shortcomings that Phoenix notes characterize studies on teenage childbearing. There is no comparison group. Sixteen through 19 year olds and women with different family backgrounds are lumped together. The sample is too small for complex statistical analyses. She does not present the British context clearly, relying on research from the United States instead. Through these design features, she has helped reify the notion that teenage mothers should be thought of as a distinct, monolithic group, a notion she rhetorically opposes. And, because of these methodological limitations, she is in no better position to make causal inferences than those who preceded her and whom she critiques. She cannot really answer the question she sets for herself, although she does provide suggestive, impressionistic evidence in support of the answer she gives it. She does a more creditable job with the implicit question. Any open-minded reader would now have doubts about conventional constructions of the "problem" of adolescent pregnancy; any such reader who is seriously interested in this issue would also now have a set of new empirical questions to test.

Phoenix makes the comparison between her study and that of Furstenberg et al. (*Adolescent Mothers in Later Life* [Cambridge University Press, 1987]). Both follow a group of teenage mothers and their children longitudinally, although Phoenix's study period is much shorter. Both provide descriptive evidence that stereotypes of teen mothers are simplistic and the consequences of early childbearing exaggerated. They diverge where Phoenix concludes that such findings imply that teenage motherhood does not contribute to disadvantage, net of the factors that select

women into teenage motherhood, while Furstenberg et al. conclude that the stereotype may be exaggerated but is not qualitatively wrong. It would be easy to cite methodological differences between the two studies (different countries, different time frames) to reconcile these opposed conclusions. But it may be more interesting, and accurate, to consider that, since neither study has an adequate comparison group, neither is designed to provide adequate scientific support for these causal inferences. The different conclusions may reflect the prior suppositions of the authors instead of conclusions made on the basis of strong empirical evidence.

Young Mothers makes a credible case against accepting common views on teen motherhood, prima facie. It is highly readable and could contribute to the education of undergraduates and graduate students. It provides interesting firsthand evidence on the experiences of young mothers, and helps place those experiences in class context. It is neither the most incisive nor lucid exposition of arguments challenging the social construction of teen motherhood as a social problem, nor does it provide adequate empirical support for the case Phoenix attempts to make. It adds to a growing body of evidence calling into question the social construction of teen motherhood as the *cause* of the social problems with which it is associated. Given the importance of what is at stake (the health and well-being of disadvantaged children and families), another attempt to stimulate open-minded thinking and spur improved empirical inquiry in this area is welcome.

▼ Basic *Sociology* ▼

CULTURE WARS
The Struggle to Define America
JAMES DAVISON HUNTER

"An extraordinary intellectual achievement—a careful and immensely constructive analysis of the sources of the moral and cultural conflicts which continue to confront us in late 20th century America."
—Robert Coles
$25.00
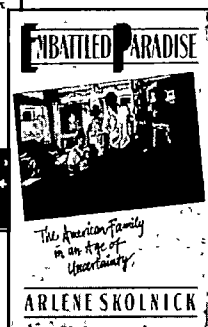
EMBATTLED PARADISE
The American Family in an Age of Uncertainty
ARLENE SKOLNICK

"At last, a reasoned and knowledgeable perspective on the American family...the clearest and most thoughtful an informed reader could hope to have."
—Cynthia Fuchs Epstein, author of *Deceptive Distinction: Sex, Gender, and the Social Order*
$23.00

FREEDOM, VOL. I
Freedom in the Making of Western Culture
ORLANDO PATTERSON

"No small pursuits for Orlando Patterson! Having tracked slavery across history, Patterson has now followed the trail of its complement, freedom...The rest of us follow breathless."
—Charles Tillly, author of *Big Structures, Large Processes, Huge Comparisons*
$29.95

NOW IN PAPERBACK!

Brave New Families
Stories of Domestic Upheaval in Late Twentieth Century America
JUDITH STACEY
$13.00

The Genocidal Mentality
Nazi Holocaust and Nuclear Threat
ROBERT JAY LIFTON
& ERIC MARKUSEN
$13.00

Indefensible Weapons
The Political and Psychological Case Against Nuclearism
Updated & Expanded Edition
ROBERT JAY LIFTON
& RICHARD FALK
$15.00

Charity Begins at Home
Generosity and Self-Interest Among the Philanthropic Elite
TERESA ODENDAHL
$13.00

Princeton

ROBERT WUTHNOW
A·C·T·S OF COMPASSION
Caring for Others and Helping Ourselves

Critics say that Americans are selfish. But statistics show that 80 million Americans—45 percent of the adult public—do volunteer work. How do we reconcile these two facets of our culture? To discover the answer, Robert Wuthnow has culled stacks of biographies, sermons, and newspaper and magazine articles; conducted a national survey; and carried out numerous in-depth interviews.

"Vigorously argued, beautifully written. . . sheds a great deal of light on a central area of American life. . . . "
—*Robert N. Bellah, coauthor of Habits of the Heart*
Cloth: $24.95 ISBN 0-691-07390-2

BETWEEN STATES AND MARKETS
The Voluntary Sector in Comparative Perspective

ROBERT WUTHNOW, EDITOR

Contributors

Helmut K. Anheier · James A. Beckford
John Boli · Helen Harding
Eliezer D. Jaffe · Michèle Lamont
Jeff Perlmutter · John Vogel
David Harrington Watt · Robert Wuthnow

Most social commentators and politicians have espoused the Tocquevillian conviction that voluntary efforts contribute vitally to the health of American society, making up a third sector "between states and markets." The essays here analyze the voluntary sector in the United States, and in seven other advanced industrial societies, to discover whether the role of voluntary associations is distinctly American. The contributors also look at a variety of related questions. Is America's voluntary sector still viable? Is it endangered by growing pressures from the government and the marketplace? And is a strong independent voluntary sector essential for creating a vibrant public sphere in which public values can be articulated?
Paper: $16.95 ISBN 0-691-02861-3
Cloth: $55.00 ISBN 0-691-09462-4

At the height of the religious ferment of the 1970s, David Van Zandt studied firsthand the most vilified of the new radical religious movements—the Children of God, or the Family of Love. First feigning membership and later gaining the permission of the Family, the author lived full-time in COG colonies in England and the Netherlands. From that experience, he has produced an informed, insightful, and humane report on how COG members function in what seems at first to be a completely bizarre setting.

This is a vivid eyewitness account. Van Zandt's frank reflections on his near-conversion experience and on the ethics of his covert observation will enrich future research with such groups.

Cloth: $24.95 ISBN 0-691-09463-2

LIVING IN THE CHILDREN OF GOD

DAVID E. VAN ZANDT
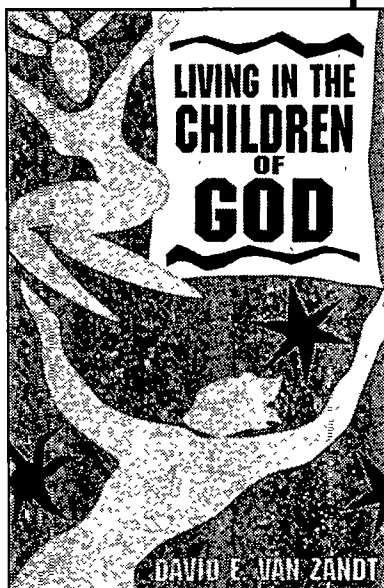
In this comparative survey of guerrilla movements in Latin America, Timothy Wickham-Crowley explores the origins and outcomes of rural insurgencies in nearly a dozen cases from 1956 onward. He explains why guerrillas have emerged strongly in certain countries but not in others.

Drawing on such materials as guerrillas' memoirs and interviews, census data and surveys on social conditions in the countryside, and national political analyses, he demonstrates that the study of revolution in Latin America must combine society-centered analyses of the sources of insurgency with state-centered analyses of the strengths or weaknesses of the regimes being confronted.

Cloth: $59.50 ISBN 0-691-07885-8

GUERRILLAS & REVOLUTION IN LATIN AMERICA

A COMPARATIVE STUDY OF INSURGENTS AND REGIMES SINCE 1956

TIMOTHY P. WICKHAM-CROWLEY
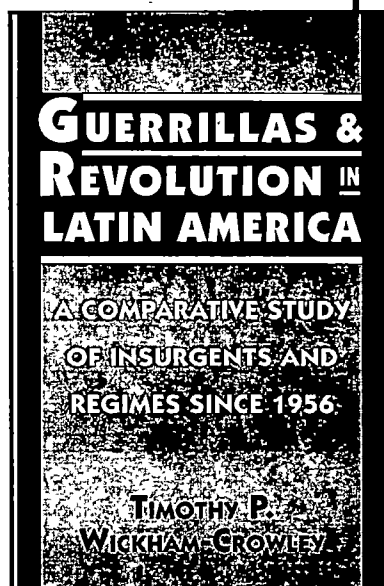
Princeton University Press

41 WILLIAM ST. • PRINCETON, NJ 08540 • (609) 258-4900
ORDERS: 800-PRS-ISBN (777-4726) • OR FROM YOUR LOCAL BOOKSTORE

ST. MARTIN'S PRESS

THE ELEMENTS OF SOCIAL SCIENTIFIC THINKING

Fifth Edition

KENNETH R. HOOVER

A concise introduction to the fundamental concepts. The fifth edition includes a discussion of the relevance of studying social science research, the political uses and abuses of social science, and a brief section on computers and statistics.

Paperbound. 190 pp. Just Published

RESEARCH METHODS IN THE SOCIAL SCIENCES

Fourth Edition

CHAVA FRANKFORT—NACHMIAS and DAVID NACHMIAS

This comprehensive text integrates the four essential components of the research process: research theory, design, data collection, and data analysis. The new edition includes recent research study examples, updated text discussions, chapter previews — to focus the student on the material — and an expanded reading list. A new appendix has been added for SPSS and for SAS.

Paperbound. 576 pp. Just Published
Instructor's Manual, Study Guide, and Computer Software available

SOCIAL INTERACTION

Readings in Sociology

Fourth Edition

CANDACE CLARK and HOWARD ROBBOY

An anthology of 40 articles chosen to teach students students from a social interactionist point of view. The articles — 13 of which are new to the fourth edition — include both classical sociological documents (Berger's "Sociology as a Form of Consciousness") and up-to-the-minute sociological debates and issues (Conrad's "The Social Meaning of AIDS").

Paperbound. 608 pp. (probable).
Publication: January 1992
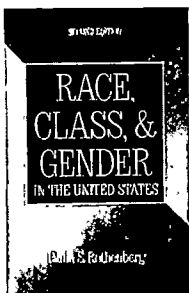
RACE, CLASS, AND GENDER IN THE UNITED STATES

Second Edition
PAULA S. ROTHENBERG

This interdisciplinary text-anthology examines ways in which racism, sexism, and — especially in the second edition— class privilege operate in contemporary U.S. society. The book includes 86 historical and legal documents, poems, statistics, pieces of fiction and nonfiction, and scholarly essays that explore the enormous complexity of interactions among race, gender, sexuality, and class.

Paperbound. 416 pp. (probable).
Publication: January 1992

SOCIETY AND TECHNOLOGICAL CHANGE

Second Edition
RUDI VOLTI

A nontechnical narrative text that describes and analyzes the causes and consequences of technological advance. The book uses timely, internationally focused examples to illustrate the social political, economic, and cultural dimensions of technology. It focuses on three social phenomena in particular — work, communications, and warfare — to help readers understand the forces that produce technological change and the consequences that result from such change.

Paperbound. 304 pp. (probable).
Publication: January 1992

FACULTY POSITION:
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Urban Studies and Planning

I POSITION. A tenured or tenure-track position in Urban/Regional Planning and Economic Development.

II JOB DESCRIPTION. The Department of Urban Studies and Planning at MIT is seeking to fill either a tenured or tenure-track position. The appointment is to teach and conduct research on economic and industrial development, and/or employment poverty and income distribution, with particular emphasis on community, urban, and regional areas within nations. The applicant's approach to these issues should emphasize planning, institutions, and public policy.

The applicant's interests must include a strong focus on the United States, and may include research on other countries.

III MINIMUM QUALIFICATIONS. The applicant should have a Ph D. in a professional field (e g , Planning, Public Policy, Business Administration) or a relevant social science discipline (e.g., Economics, etc.). Several years of research experience is essential as well as the ability to teach students from diverse academic and cultural backgrounds.

IV. APPLICATION The deadline for applicants is January 31, 1992. Applicants should include a statement of their teaching and research interests as well as C.V, and will be considered as they are received. Early submissions are encouraged and should be sent to:

> Development Search Committee
> 9-519, M I.T.
> Cambridge, Mass. 02139

AJS

The University of Chicago Press

CONTENTS

Review Essay

Book Reviews

IN THIS ISSUE

DOUG MCADAM, professor of sociology at the University of Arizona, is currently a fellow at the Center for Advanced Study in the Behavioral Sciences at Stanford University. He is the author of *Political Process and the Development of Black Insurgency* and *Freedom Summer* (the latter is a co-winner of the 1990 C. Wright Mills Award). His current research interests include the dynamics of strategic action, the diffusion of protest behavior, demography and social movements, and meaning and membership in social life.

WILLIAM T. BIELBY is professor of sociology at the University of California, Santa Barbara. His research interests include organizations, popular culture, gender, and quantitative methods. His current research examines how changing organizational arrangements and business conditions shape decision making among network programmers and the producers of television series.

DENISE D. BIELBY is associate professor of sociology at the University of California, Santa Barbara. Her research interests are in the areas of gender, work, and popular culture. Her current research examines social networks and careers in the television industry, most recently focusing on the coproduction of meaning between viewers and producers of television series.

GARY ALAN FINE is professor and head of the Department of Sociology at the University of Georgia. He is currently completing a book entitled *Kitchens: An Interpretivist Approach to Organizational Culture and Structure,* which examines the hospitality industry in light of occupational rhetoric, aesthetic talk, and the connection of small group culture and the economic order. He is currently conducting research on the social world of high school debate.

EMANUEL A. SCHEGLOFF is professor of sociology at the University of California, Los Angeles. His work is taken up with the naturalistic study of talk and other forms of conduct in interaction, both for their bearing on traditional concerns of social science and for the new problems and concerns such inquiry allows us, and forces us, to formulate and address.

JOACHIM J. SAVELSBERG is assistant professor of sociology at the University of Minnesota. His present research is on the emergence and diffusion of crime and punishment related knowledge and its impact on legal and political decision making. Previous books and articles deal with economic crime legislation, sentencing, and criminology in West Germany. A former research fellow at Johns Hopkins and Harvard universities, Savelsberg recently moved to the United States.

DEREK SAYER is professor of sociology at the University of Alberta. His most recent book is *Capitalism and Modernity: An Excursus on Marx and Weber*. Currently he is working at the Institute of Sociology of the Czechoslovak Academy of Sciences, on the history of Prague, 1848–1990. He is a fellow of the Royal Historical Society and a founding editor of the *Journal of Historical Sociology*.

FRANK DOBBIN is assistant professor at Princeton University. He studies the emergence of public policy paradigms and has recently completed a book on the development of national industrial policy strategies in the United States, France, and Britain. He also studies the effect of state institutions on the evolution of organizational practices and is currently analyzing changing entrepreneurial strategies among early railway firms (with Timothy Dowd), shifts in personnel practices since the 1950s (with John Meyer, John Sutton, and W. Richard Scott), and cross-national differences in management practices (with Terry Boychuk).

JOHN F. PADGETT is associate professor of political science at the University of Chicago. His current research project, based on extensive primary sources, is a network analysis of political party formation in Renaissance Florence.

Information for Contributors

Editorial Procedures: All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript. Attach a cover page with the title of the manuscript, your name, and your affiliation. All identifying references and footnotes should appear on a separate page. A legible, carefully prepared manuscript will facilitate the work of readers. Manuscripts are accepted for publication subject to nonsubstantive editing with the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his article without charge in any book of which he is the author or editor after it has appeared in the *Journal*.

Manuscript Acceptance Policy: While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the journal has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the journal in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double-spaced,* allowing generous margins at top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Avoid proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, including acknowledgment footnote (which should be "1"). *Provide a word count on the cover page of the manuscript.*

2. Type each table on a separate page. Refer to each table in numerical order in the text. Prepare tables without vertical lines. Tables should not contain more than 20 two-digit columns or the equivalent.

3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Draw figures on white paper with india ink. Original or glossy print of professionally drawn figure will be required if manuscript is accepted.

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6. *Five copies* of the manuscript and abstract, on white paper, must be submitted. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, The University of Chicago, Chicago, Illinois 60637 *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard, containing the manuscript title, to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or money order for $15.00, payable to The University of Chicago Press in U.S. currency or its equivalent by money order or bank draft. Papers will not be reviewed until this fee has been paid The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

Format of References in Text: All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963)

2. Pagination follows year of publication: (Lipset 1964, pp. 61–65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of Census 1963, p 117).

4 With more than one reference to an author in the same year, distinguish them by use of letters (*a, b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons

Format of References: List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "References." Examples follow

Davis, James Allan. 1978. *General Social Survey, 1972–1978: Cumulative Data* (MRDF). NORC ed. Chicago· NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

Davis, K. 1963*a*. "Social Demography." Pp. 124–37 in *The Behavioral Sciences Today,* edited by Bernard Berelson. New York. Basic.

———. 1963*b*. "The Theory of Change and Response in Modern Demographic History." *Population Index* 29:345–66.

Goode, W. J. 1967. "The Protection of the Inept." *American Sociological Review* 32:5–19.

Moore, Wilbert E., and Arnold S. Feldman. 1960. *Labor Commitment and Social Change in Developing Areas.* New York: Social Science Research Council.

Sanford, Nevitt, ed. 1962. *The American College.* New York: Wiley.

U.S. Bureau of the Census. 1970. *Census of Population and Housing 1970, Summary Statistic File 4H: U.S.* (MRDF). DUALabs ed. Washington, D.C.: Bureau of the Census. Distributed by DUALabs, Rosslyn, Va.

Weber, Max. (1921) 1968. "Society's Problems." Pp. 12–16 in *Economy and Society,* edited by Guenther Roth and Claus Wittich. New York: Bedminister.

Gender as a Mediator of the Activist Experience: The Case of Freedom Summer[1]

Doug McAdam
University of Arizona

Using data on 330 applicants to the 1964 Mississippi Freedom Summer project, the author seeks to assess the effect of gender on all phases of the activist process. The results indicate that gender powerfully influenced the dynamics of recruitment to the project, experiences during the summer, the long-term political effects that followed from participation in the campaign, and the subjects' own assessment of the project's impact on their lives. Interestingly, the behavioral effects of the project are greatest for the male volunteers. But it is the female volunteers who *attribute* the greatest personal significance to the project. It is suggested that this seeming paradox can be explained by two factors: (1) the more extensive histories of activism the female volunteers brought to the project and (2) the significance assigned to the project in feminist accounts of the origins of the women's liberation movement.

In recent years the topic studied most frequently by scholars interested in the dynamics of individual activism is that of "differential recruitment." Why does one person participate in a social movement while another fails to do so? Underlying this question is the presumption that there exists some finite set of factors that differentiates activists from nonactivists and therefore explains movement participation. Many individual factors have been proposed as explanations for activism: "authoritarian personality characteristics" (Hoffer 1951; Lipset and Rabb 1973),

social isolation and alienation (Aberle 1966; Kornhauser 1959), feelings of "relative deprivation" (Davies 1969; Feierabend, Feierabend, and Nesvold 1969; Geschwender 1968; Gurr 1970; Morrison 1973), consistency between movement goals and parental values (Block 1972; Flacks 1967; Keniston 1968), selective incentives for participation (Oliver 1984; Oliver, Marwell, and Teixeira 1985), prior contact with a member of the movement (Briet, Klandermans, and Kroon 1984; Gerlach and Hine 1970; Heirich 1977; McAdam 1986; Orum 1972, Snow, Zurcher, and Ekland-Olson 1980), integration into activist networks (Fernandez and McAdam 1988; Rosenthal et al. 1985), and the absence of "biographical constraints" that might otherwise discourage a person from getting involved in a movement (McAdam 1986; McCarthy and Zald 1973; Snow and Rochford 1983, pp. 2–3).

The diversity and sheer number of factors proposed as explanations of activism attest to the considerable interest and controversy the recruitment question has generated among movement scholars. For all of the differences among these approaches, however, there remains a single underlying similarity. All of these studies are premised on the idea that activists can be distinguished from nonactivists on the basis of some factor or set of factors. Based on the large number of positive relationships these studies have produced, this assumption would seem to have considerable merit. Activists do seem to be distinguishable from nonactivists. It is less certain that activists constitute a homogeneous population in their own right. Yet very little research attention has been paid to differences between activists within the same movement. Researchers have been so concerned with differentiating activists from nonactivists that they have implicitly depicted movement participants as cut from whole cloth. This characterization would seem to be unwarranted. As social phenomena, social movements are large and diverse enough to provide a broad umbrella under which a wide variety of participants can huddle. It should, therefore, be just as interesting to study variation among participants in a single movement, as it is to delineate differences between activists and nonactivists.

Indeed, the few empirical studies that have taken this approach have shown just how fruitful it can be. Fendrich (1977), for example, was able to demonstrate that the biographical consequences of participation in the civil rights movement were very different for black and white activists. In their study of participants in the sanctuary movement, Wiltfang and McAdam (1991) have shown that activists can be distinguished on the basis of the levels of "cost" and "risk" associated with their activities (see also McAdam 1986). These differences, in turn, would seem to reflect variation in recruitment dynamics as well as in the consequences of movement participation.

In this article, however, I want to focus on yet another—and arguably the most important—factor mediating not only the activist experience, but all of social life. I am referring to gender. Specifically, I am interested in the many and complex ways in which gender mediates the activist experience. Recent advances in feminist theory (e.g., Connell 1987; Scott 1988) have refined our understanding of the ways in which gender may shape lived experience. It is not simply a matter of socialized differences in behavior and attitudes between men and women. Elaborate belief systems about maleness and femaleness get encoded in social structure in myriad ways. The result is a complex system of constraints and opportunities that powerfully shapes the experiences that men and women can have, what they can know, and who they can become. Moreover, even within a single society, gender is always experienced within equally elaborated systems of race, class, and age. The unique intersection of these various systems, then, further differentiates the lived experience of gender.[2] Much of the conceptual richness and complexity reflected in these newer perspectives on gender has been incorporated into the growing body of feminist scholarship on women's participation in social movements. Unfortunately, the empirical literature on individual activism has been only lightly touched by this conceptual revolution.

GENDER AS A FOCUS OF SOCIAL MOVEMENT STUDY

The study of women's participation in social movements has grown apace of the general rise of feminist scholarship in the social sciences, so that now there exists a rich and varied literature on the topic. Among the earliest works in the area were several ground-breaking studies of the emergence and subsequent development of the modern women's movement (Carden 1974; Freeman 1973, 1975; Hole and Levine 1971; Morgan 1970). The development of an academic feminism also sparked a renewed interest in the earlier women's rights and suffrage movements (Chafetz and Dworkin 1986; O'Neill 1969).

Since then, however, a wide range of specialized literatures on the general topic of women and activism has emerged. For example, a number of historians and sociologists have sought to describe and account for

---

[2] Unfortunately, the homogeneity of my subject population precludes any serious attention to race, class, age, and gender interactions My subjects were drawn from among all those who made formal application to the 1964 Mississippi Freedom Summer Project. Overwhelmingly, these were middle-class and upper-middle-class, white college students (McAdam 1988, pp. 41–44). Accordingly, there was very little class, race, or age variation among my subjects that might have been used as the empirical basis for an investigation of the complex ways in which race, class, age, *and* gender combined to shape the Freedom Summer experience.

women's involvement in the labor movement (see Dye 1980; Foner 1979; Kessler-Harris 1985; Milkman 1985, 1980; Wertheimer 1977). Others have analyzed other forms of grass-roots organizing by working- and lower-class women (Bookman and Morgen 1988; Canterow 1980; Hertz 1977; McCourt 1977; Sacks 1988; West 1981). Still others have sought to disentangle the complex interplay of class, race, and gender among female activists (Aptheker 1982; Blumberg and Roye 1979; Blumberg 1980; Dill 1981; Luttrell 1988). Finally, a growing number of scholars have studied the involvement of women in the New Right (see Brady and Tedin 1976; Dworkin 1983; Klatch 1987, 1988; Luker 1984).

But for all the richness and diversity of these literatures, they have had little impact on the sociological study of individual activism.[3] Indeed, the sociological study of social movements remains only lightly touched by the "gender revolution" in the social sciences.[4] What Millman and Kanter (1975) wrote nearly 15 years ago about sociology as a whole, and Stacey and Thorne (1985) echoed 10 years later, still applies to the sociological study of individual activism. "Sociology," according to Millman and Kanter, "often assumes a 'single society' with respect to men and women, in which generalizations can be made about all participants, yet men and women may inhabit different social words, and these must be taken into account" (1975, p. xiii). Since Millman and Kanter penned this line, scholars have produced compelling evidence to suggest that males and females may experience such diverse phenomena as marriage (Bernard 1972), divorce (Weitzman 1985), justice (Gilligan 1972, 1977), emotions (Hochschild 1983), and politics (McCormack 1975; Sapiro 1983) in significantly different ways. Why should the activist experience be any different? Indeed, given male/female differences in socialization and the powerful ways that differences in gender-based behavioral expectations and opportunities (see Gerson 1985) get embedded in social structure,

---

[3] This state of affairs owes, in part, to narrow subfield boundaries and the failure of movement scholars to mine the relevant feminist scholarship for the insights it would afford. At the same time, it is important to remember that these two literatures have quite different goals, which make them less relevant to one another than it would appear at first blush. Feminist scholars seek to understand the ways in which the structures and circumstances of women's lives afford them the opportunity and motive to engage in collective action. Movement scholars, on the other hand, are concerned with teasing out the generic dynamics of individual activism. I am concerned with the latter agenda, as well, but want to broaden it to include the mediating affects of gender and a host of other variables

[4] To my knowledge, only Rochford (1985), in a chapter in his book on the Hare Krishna, Lawson and Barton (1980), in an article on the tenant movement in New York City, and Cable (1992), in an interesting paper on women's involvement in a local environmental movement, have sought to analyze the effects of gender on activism.

why should we presume that males and females experience *any* social phenomenon in the same way? It would be better to presume differences that can then be explored empirically, and that is the tack that I will take in this article.

THE STUDY

In seeking to analyze the ways in which gender mediates movement participation, I will focus on a single instance of high-risk activism: participation in the 1964 Mississippi Freedom Summer project. That campaign brought hundreds of primarily white, Northern college students to Mississippi for all, or part of, the summer of 1964 to help staff "freedom schools," register black voters, and dramatize the continued denial of civil rights throughout the South. As instances of activism go, the summer project was time-consuming, physically demanding, and highly newsworthy.

The project itself began in early June with the first contingent of volunteers arriving in Mississippi fresh from a week of training at Oxford, Ohio. Within 10 days, three project members, Mickey Schwerner, James Chaney, and Andrew Goodman, were kidnapped and killed by a group of segregationists led by Mississippi law enforcement officers. That event set the tone for the summer as the remaining volunteers endured beatings, bombings, and arrests. Moreover, most did so while sharing the grinding poverty and unrelieved tension that was the daily lot of the black families that housed them.

Preliminary to their participation in the campaign, all prospective volunteers filled out detailed applications providing information on, among other topics, their organizational affiliations, previous civil rights activities, and reasons for volunteering. On the basis of these applications (and, on occasion, subsequent interviews), the prospective volunteer was either accepted or rejected. Acceptance did not necessarily mean participation in the campaign, however. In advance of the summer, many of the accepted applicants informed campaign staffers that they would not be taking part in the summer effort after all. Completed applications for all three groups—rejects, participants, and withdrawals—were copied from the originals, which are now housed in the archives of the Martin Luther King, Jr., Center for the Study of Nonviolence in Atlanta and in the New Mississippi Foundation in Jackson, Mississippi.[5] A total of 1,068

applications were coded in connection with this study. The breakdown of these applications shows 720 participants, 239 withdrawals, 55 rejections, and 54 whose status in the summer project is unclear. Together these applications provide a unique source of data for assessing the relative importance of various factors—including gender—in shaping *recruitment* to activism.

But to fully explore the question of gender and activism, we will want to examine the volunteers' summer experiences and subsequent biographies as well. To do so I made use of survey and interview data gathered in 1983–84 as part of an exhaustive follow-up study of all Freedom Summer applicants. That study used the original project applications as its methodological point of departure. Specifically, several items of information from the original applications—alma mater, parents' address, major in school—functioned as crucial leads in my efforts to obtain current addresses for as many of the applicants as possible.

The results of these efforts were verified current addresses for 556 of the 959 participants and withdrawals for whom I had applications. Of these, 382 (of a total of 720) had been participants in the project, while another 174 (of 239) had withdrawn in advance of the summer. Separate questionnaires were then prepared and sent to the participants and to the "no-shows." Participants were questioned about their experiences during Freedom Summer, their activist histories, and the broad contours of their lives, personal as well as political, after Freedom Summer. The questionnaire sent to the no-shows dealt with these latter two topics. In all, 212 (or 56%) of the participants—68 females and 144 males—and 118 (or 68%) of the no-shows—50 females and 68 males—returned completed questionnaires. In addition, in-depth interviews—oral history sessions, really—were conducted with 40 volunteers and another 40 no-shows to flesh out the information gleaned from the questionnaires. These interviews proved invaluable in providing insights into the complex ways in which gender shaped the volunteers' behaviors and perceptions both during and in the years since the summer. My use of oral histories is consistent with a growing trend in the application of the technique to the study of social movements. The work of scholars such as Ron Fraser (1988) on the international student movement of the 1960s, Howell Raines (1983) on the civil rights movement, and Rebecca Klatch (1987) on the women of the New Right, attests to the value of the technique.

RESULTS

In seeking to determine how gender mediated participation in the Freedom Summer project, I will look at three potential sources of difference between males and females: (1) differences in the recruitment of men

TABLE 1

STATUS ON THE SUMMER PROJECT BY GENDER (%)

Gender	All Applicants	Rejects	No-Shows	Volunteers
Male	59	48	52	62
Female	41	52	48	38

and women to the project, (2) differences in their experiences during the summer, and (3) differences in the political consequences that followed from involvement in the project. I will take up these sources of difference chronologically, beginning with recruitment to the project.

Recruitment to the Project

In a previous analysis of recruitment to Freedom Summer (McAdam 1986), gender was found to be one of six variables that were significantly related to a person's chances of participating in the project.[6] Specifically, female applicants were less likely to make it to Mississippi than were their male counterparts. Table 1 shows that the female applicants were disproportionately concentrated in the ranks of those who were rejected for participation or who, despite being accepted, failed to take part in the project.

   It is the female applicants' lower likelihood of participation that begs for explanation. Two classes of possible explanations can be distinguished. One would focus on characteristics of the female applicants as the key to their lower rates of participation in the project. Perhaps the female applicants simply possessed fewer of the characteristics found in the previous study to be associated with involvement in Freedom Summer. The other possibility is that the application and recruitment process posed greater obstacles to female than to male participation in the project.

   The logical place to begin an analysis of the first explanation is with the remaining individual factors previously shown to predict involvement

[6] The analysis in question was a logistic regression assessing the effects of various independent variables on the applicant's likelihood of participation in the Freedom Summer project. Besides the six significant variables, the other variables included in the analysis were the sum of the subject's "biographical constraints" (e g., married, full-time employment, etc.) on the eve of the project, their major in college, their home region, the geographic region of their college, their race, highest grade they completed, and the distance from the subject's home to Mississippi. The full analysis is reported elsewhere (McAdam 1986, p. 86).

1217

TABLE 2

SIGNIFICANT PREDICTORS OF FREEDOM SUMMER PARTICIPANTS
AND THE DIRECTION OF ASSOCIATION

Predictor	Direction of Association
No. of organizational affiliations	Positive
Level of prior civil rights activism . ..	Positive
Strong tie to another volunteer (if yes) .	Positive
Strong tie to another no-show (if yes)..	Negative
Age	Positive

in the summer project. These five factors and the direction of their association with Freedom Summer participation are given in table 2.[7]

In effect, these five variables constitute the individual "qualifications" for participation in the summer project. On the basis of these characteristics, is it possible that the female applicants were somehow "less qualified" for activism than their male counterparts? Table 3 clearly suggests otherwise.

On two of the variables—level of civil rights activism and number of organizational affiliations—there were significant differences by gender. And, in both cases, it was the female rather than male applicants who emerged as "more qualified" for participation in the project. Indeed, the female applicants ranked higher than their male counterparts on four of the five "qualification" variables. So clearly there was nothing in the

[7] Several of the variables referred to in the text require some further discussion. The two "strong-tie" variables were constructed using information drawn from a single item on the original project application. The item asked the subject to list at least 10 persons whom they wished kept informed of their summer activities. Most often, the names supplied by the applicants were those of parents, parents' friends, professors, ministers, and other noteworthy or influential adults they had contact with. Quite often, however, applicants would list another applicant. This enabled me to construct a measure of the interpersonal ties connecting participants and withdrawls to (a) other Freedom Summer volunteers and (b) withdrawals from the project. In doing so, I was careful to distinguish between "strong" and "weak" ties (Granovetter 1973). Persons listed directly on the subject's application were designated as strong ties. Weak ties were defined as persons who, although they were not listed on the subject's application, were nonetheless linked by way of an intervening strong tie. The other variable that requires some explanation is the measure of prior civil rights activism Both the participants and the no-shows were asked to list on their applications all previous involvements in civil rights activities. In coding these activities, I assigned a numerical value to each that reflected its intensity relative to all other forms of civil rights activism. So, e.g., participation in the Freedom Rides was assigned a score of "7," while contributing money to a civil rights organization had a designated point value of only "1." I then gave each subject a final activity score that was the sum of the point totals for the activities reported on their applications.

TABLE 3

SELECTED CHARACTERISTICS OF THE FREEDOM SUMMER APPLICANTS BY GENDER

	Males	Females
Mean no. of organizational affiliations	2.16	2.39*
Level of prior civil rights activism	4 72	5.59*
Those with strong tie to another volunteer (%)..........	38	34
Those with strong tie to another no-show (%)	8	7
Mean age	23.00	23.46

* $P < .05$.

individual profile of the female applicants that can account for their underrepresentation in the ranks of those who actually made it to Mississippi. This encourages us to examine the broader context in which males and females applied to the Freedom Summer project.

Perhaps the most relevant contextual factor to keep in mind is the generally traditional conception of men and women still prevalent in the United States on the eve of Freedom Summer. As commentators from Betty Friedan (1963) to Sylvia Hewlett (1986) remind us, the period encompassing the 1950s and early 1960s was one of the most conservative eras of gender socialization in U.S. history. Given this context, it is surprising that so many females applied to the project. "Certainly, the dangerous, unchaperoned, interracial, and political character of the project posed more of a threat to [the then current] conceptions of femaleness than maleness. For the male applicant, participation in a project like Freedom Summer could be seen as functionally equivalent to any number of other traditional challenges that were available to young men as part of the process of 'becoming a man,' 'seeing the world,' or 'testing one's mettle'" (McAdam 1988, p. 57). There were few, if any, equivalent precedents available to the female volunteers to legitimate their participation in the project. Unlike young men, college-age women were not expected to expose themselves to adventures of a potentially dangerous nature.

The traditional sexual double standard also served as an impediment to female involvement in the Freedom Summer project. The unchaperoned nature of the project was much more likely to be of concern to the parents of female as opposed to male applicants. Finally, the interracial character of the project raised the specter of that most potent taboo—contact between white females and black males—and served as yet another serious obstacle to female, but not male, involvement in the summer project.

In view of archival evidence and my interviews with 80 applicants, it seems clear that these various factors did indeed constitute a set of barri-

,ers to female participation in the project. Especially potent was the combination of the sexual double standard and the taboo concerning contact between white women and black men. One interesting form of archival evidence helps to illustrate this point. As part of the application process, many of the prospective volunteers were interviewed by a member of the project staff or an approved interviewer at the applicant's school. In virtually all cases, the interviewer was white, with women being slightly more likely than men to serve in this capacity.[8] Following the interview, the interviewer was supposed to forward a written report to the central project office summarizing their evaluation of the applicant. A good many of the "interviewer reports" from these sessions have been obtained from a variety of collections around the country.[9] They are important for what they reveal about the way project staff viewed the applicants on the eve of the summer. On the issue of gender differences, these reports support one very simple conclusion: the substance and tone of the interviewer's comments differ depending on the sex of the applicant. "One difference concerns the amount of attention given the issue of sex in the interviews. Rarely, if ever, does the issue seem to have been discussed in the interviews with male applicants. Quite the opposite is true in regards to the sessions with female applicants" (McAdam 1988, pp. 58–59). One interviewer report form contains the following exchange between an interviewer and a female applicant:

> A ___ B ___ was in on the interview and asked her how she'd deal with a Negro man who came up to her on the street and asked to sleep with her. She said she might. I asked her what she'd do if I told her, as a staff member, that sexual activity would endanger everyone, and not to do it. She said she might go ahead and do it anyhow.

The applicant in question was rejected for work on the project. Elsewhere I have discussed the logic and possible influences on such decisions.

> Given the very real Southern phobia regarding contact between black males and white females as well as the need for discipline on the project, this seems like a reasonable decision. The more basic question is, why were these issues more often raised with female applicants? Was it because the issue of interracial sexuality was more explosive when it came to white

.

[8] Of the 36 interviewer reports I obtained, the race of the interviewer is known in 34 cases. In 30 of these 34 reports, the interviewer is white. The sex of the interviewer is known in 35 of the cases, with women serving as interviewers in 21 of these 35 interviews.

[9] Those I have obtained came from three sources: the archives of the Martin Luther King, Jr., Center in Atlanta, Georgia; the Freedom Information Service in Jackson, Mississippi; and the State Historical Society of Wisconsin Archives in Madison.

women than white men? Or does the different emphasis in the interviews perhaps also reflect the more general influence of a traditional sexual double standard that made sex less acceptable for a woman than a man? There is no clear answer to this question. What is certain is that the whole issue of sexuality represented a barrier to female but not male participation in the project. [McAdam 1988, p 59]

The frequent references to the physical appearance of the female applicants provided yet another example of the influence of a sexual double standard on the interview process. The following two excerpts typify the substance the tone of these references: "This girl is uncommonly ugly with an almost piggish look to her. EEEECH!" and "She has long blond hair and is frankly a knock-out."[10]

Other than a single reference to the "frailness" of one of the prospective male volunteers, the interviewer reports on the male applicants make no mention of physical appearance. Although there is no evidence that female applicants were rejected on the basis of their physical appearance, the comments suggest that the female applicants were being delivered a contradictory set of signals about the issue of sexuality on the summer project. At the same time they were being warned of the dangers of sex in Mississippi, they were also being physically evaluated in a way their male counterparts were not.

Finally, the following interviewer report hints at a kind of generalized sexism that was also to show itself in Mississippi:

> This girl presents problems. She is a photographer; graduate of Radcliffe '61, works doing odds and ends as she takes pictures. Is interested in doing documentary films as a profession. She is a trained typist, but hates to do this type of work. Has had some experience in civil rights, BUT didn't "like" working for NAACP because she walked in to help and found fifteen girls typing and the room filled with smoke. . . . The girl flatly refused to do clerical work. When pushed to the wall she said she'd do it, but not more than half the summer. . . . She says she wants to "work with people" and "learn something from the summer," and that she won't be able to do either if she works in an office. . . . RECOMMENDATION. *DEFINITE REJECTION.*

One of the applicants I spoke with described her interview experience as similar to that of the woman discussed above.

> Officially I'm sure they [the project staff] thought I had withdrawn from the project. But . . . that's not what happened. I've always felt that they *rejected* me. . . . I mean the whole interview was so hostile and sexist. I wouldn't have known to call it that at the time, but that's certainly what was going on. I remember the gal who was interviewing me asking if I'd be willing to do clerical work. I said, "Sure, but I'd like to work on voter

---

[10] Interviewer report form.

> registration too." At that point she launches into this harangue about how "we need dutiful foot soldiers," and "you've got to accept project discipline," blah, blah, blah. . . The whole tone was, "They also serve who sit and type!" I mean, I can't believe they were putting the same questions to the guys.[11]

While the *general* issues of project discipline and staff authority seem to have been a central concern in *all* of the interviews, the *specific* questions regarding acceptance of work assignments appear to have been asked of the female applicants far more often than their male counterparts. "The impression that one is left with is that the project staff had a narrower vision of women's role in the project than the one they applied to male applicants. Given the prevailing attitudes toward women and work in early Sixties America, such a vision is hardly surprising. The application of this vision, however, represented yet another obstacle to women's participation in the project" (McAdam 1988, p. 60). Nor was it only the project staff that posed an obstacle to female participation in the project. The interviews conducted with the volunteers and no-shows indicated that females were apt to encounter greater opposition from their families, ministers, and others as well. For instance, nearly half of the female no-shows (eight of 17) I interviewed attributed a great deal of importance to parental opposition in explaining their withdrawal from the project. Only six of 23 male no-shows did likewise during their interviews.

In the face of the greater opposition females encountered in applying to Freedom Summer, what enabled some of them to overcome these obstacles and make it to Mississippi? The answer would appear to be reflected in the data presented in table 4, which reports the results of a logit regression analysis predicting participation in Freedom Summer.

For the most part, the data presented in table 4 confirm findings reported elsewhere (see McAdam 1986). However, there is a significant exception to this rule. The introduction of gender interactions washes out the significant main effect of gender reported in the earlier article. Specifically, it is the interaction of gender and the number of political affiliations one had before Freedom Summer that appears to mediate the relationship between gender and project status. There are two possible interpretations of this finding. The first is that the female applicants benefited from or perhaps were more dependent on organizational support and sponsorship for their participation in the project than were males. Perhaps being a member of a campus political group, civil rights organization, or other such groups afforded the female applicants the countervailing social support needed to offset the other forms of opposition they were likely to encounter. In this view, the support—and per-

---

[11] Author's interview with an anonymous Freedom Summer applicant, June 6, 1985.

TABLE 4

RESULTS OF LOGIT REGRESSION PREDICTING PARTICIPATION IN THE FREEDOM
SUMMER PROJECT ($N = 886$)

INDEPENDENT VARIABLE[a]	DEPENDENT VARIABLE = SUMMER STATUS[b]	
	b	SE(b)
Strong tie to another volunteer	1.0781**	.330
Weak tie to another volunteer	−.132	.226
Strong tie to a no-show	− 772	.579
Race (0 = white; 1 = black)	073	.196
No of political affiliations,		
presummer.	.250**	.083
Level of prior activism	018	023
Age	.064**	024
Highest grade completed....	−.048*	020
Distance from home to Mississippi	0001	0001
Gender (0 = women, 1 = men) .. .	197	889
Gender × strong tie to volunteer	− 189	.414
Gender × strong tie to no-show . .	861	.760
Gender × age	021	037
Gender × no. of affiliations.	− 236*	.103
Gender × level of prior activism . ..	.029	.032
Constant......................	− 875	.666

[a] For details of how the independent variables were coded, see McAdam (1986, pp 82–86)
[b] Summer status 0 = no-shows; 1 = volunteer
* $P < .05$
** $P < 01$

haps active participation—of other group members might have served
to insulate the female applicants from parental or general societal criti-
cism of their actions. This is at least a plausible interpretation based on
the data reported in table 4.

The second possibility rests on a very different interpretation of the
significance of this variable. Number of organizational affiliations may
simply be a proxy for the level of a person's commitment and involve-
ment in movement politics. As such the significant positive coefficient
attached to the affilliation variable in the equation for female applicants
may well indicate a deeper commitment to the movement on the part of
some subset of the female applicants. It was this greater commitment
that enabled the female volunteers to overcome the severe opposition
they encountered. The data do not allow us to choose between these
alternative interpretations. As we will see, however, this second account
accords nicely with other findings to be reported later. Whichever inter-
pretation is correct—and both processes may be operating—it is clear

that, in the case of Freedom Summer, the dynamics of recuitment varied by gender.

Experiences in Mississippi

What about the volunteers' experiences in Mississippi? Did gender mediate those as well? The answer is clearly yes. This is not to say that there were no similarities between the male and female experiences of Freedom Summer. On the contrary, my interviews with the volunteers strongly affirm the shared nature of the experience. Irrespective of gender, the summer project was experienced as powerfully affecting and overwhelmingly positive by the vast majority of the volunteers.

In expressing this view the female volunteers tend to emphasize the same positive features of the project as their male counterparts. Perhaps most important, the volunteers, male and female alike, remember the project for the good it did their self-confidence and sense of self-worth. They had been tried and were not found wanting. As one male volunteer put it: "The memories of that summer are very important to me because they . . . redeem me personally. . . . [They serve] as a reminder to me that there are qualities in me that are worth . . . something and that people are capable of quite remarkable things. It's the single most enduring . . . moment of my life. I believe in it beyond anything."[12]

The ideology of Freedom Summer and of the Student Nonviolent Coordinating Committee (SNCC) staff who were responsible for the project's organization only reinforced the positive personal messages most of the female volunteers took from the summer. Even if its practice sometimes contradicted its ideals, SNCC nonetheless celebrated the dignity and the fundamental decision-making ability of each individual, translating these themes into calls for self-determination and participatory democracy for all peoples. These themes gave the volunteers an ideology and language readily applicable to their own lives and to the issues that would dominate the New Left in the years to come. Not surprisingly, the volunteers were among the first to invoke the ideological lessons of Mississippi in the burgeoning student, antiwar, and women's liberation movements (McAdam 1988, chap. 5). Thus besides the *individual* psychological payoffs of participation, the volunteers benefited from a kind of *collective* political education as well.

Finally, nearly all of the volunteers retain strong and positive memories of the rich, tight-knit sense of community they experienced in Mississippi. Said one female volunteer, "What I remember about Mississippi was the love I felt . . . from everyone. There was this openness and acceptance

[12] Author's interview with an anonymous Freedom Summer volunteer, April 22, 1985.

of you as a person that I've never really felt since."[13] For many of the volunteers, raised in the somewhat anomic propriety of 1950s white, middle-class America, this sense of community was and remains very appealing.

So there was a fundamental, positive continuity to the way the male and female volunteers experienced Freedom Summer. But there were significant differences of a less positive sort as well. As noted earlier, the SNCC rhetoric of equality and personal dignity was not always realized in practice. Instead, the female volunteers were exposed to forms of discrimination that, in the face of the project's more egalitarian rhetoric, began to produce in them an experiential awareness of sexism.

The two principal forms of discrimination the female volunteers experienced centered on the issues of work assignments and sexual politics. On the sexual front, women were subjected to considerable harassment and a clear double standard regarding sexual behavior. One volunteer captured the dilemma nicely when she said: "It really was your classic 'damned-if-you-do, damned-if-you-don't' situation. If you didn't [have sex], you could count on being harassed. If you did, you ran the risk of being written off as a 'bad girl' and tossed off the project. This didn't happen to the guys."[14]

In fact, a number of women *were* asked to leave the project for behavior that was considered unbecoming a project member.[15] The accounts of these incidents, however, always betray more than a hint of a sexual double standard. In a journal written during the summer, one volunteer described how "some people found it very hard to conform to the necessary discipline . . . in particular one girl was sent home last weekend after various incidents involving breaches of discipline in the field of social and public etiquette."[16] Notwithstanding the importance of maintaining project discipline in a situation as volatile as Mississippi, it remains hard to understand how a female volunteer could have violated "social and public etiquette" without the help of a male. However, I found no evidence of any male—white or black, staff or volunteer— being removed from any of the projects for sexual (or any other) reasons.

Nor was this double standard evident only in regard to the volunteers. One staff member remembers it being applied within SNCC, as well. "I

[13] This quote was taken from a transcript of a taped session of shared reminiscences offered by some 30 Bay Area volunteers gathered together for a 20-year reunion (tapes in author's possession).

[14] Interview with an anonymous Freedom Summer volunteer, March 21, 1985.

[15] Two volunteers recounted such incidents in their interviews with me (August 23, 1984, and June 12, 1985, respectively), while a third described a similar case in his journal.

[16] Entry in Ronald de Sousa's journal (in author's possession).

remember huge tensions in the SNCC staff. . . . A lot of it was on who dated who. I remember . . . all these black guys were dating the white volunteers, and then one of the black girls . . . had one date one night with a white guy. And I heard that the next morning four black . . . male SNCC staff were over at her house chewing her out."[17]

Discrimination was also evident in the way that work was structured on the project. Overwhelmingly, the women were utilized as freedom school teachers, clerical workers, or community center staff. The bulk of the voter registration work on the project was left to the males (McAdam 1988, p. 107). It is important to emphasize that these divisions reflect more than simple sex differences in socialized preferences for work assignment. Here gender was not so much a matter of differences in socialization as a concrete aspect of social structure limiting the opportunities available to the female volunteers. Table 5 provides a comparison of the preferred and actual work assignments given the male and female volunteers.

Reflecting the overall importance of the freedom schools and voter registration to the project, it is not surprising that the vast majority of volunteers—male and female alike—were concentrated in these two areas. What is significant is the marked disparity in the actual and preferred work assignments given the female volunteers. Once again, these assignments reflect a narrower and less political role for women on the project than the one available to the male volunteers. Part of the logic underlying this less political conception of women's role on the project had little to do with sexism and is entirely understandable. As I have noted elsewhere, "many on the project staff felt the mere presence of white women on the project was sufficiently threatening to the white community, that to have them canvassing door to door was to only court disaster. The overall safety of the project would be better served, it was felt, by placing the women in less visible teaching and clerical positions" (McAdam 1988, p. 108).

The net effect of this policy was to reproduce traditional gender and work roles on the projects. The men would leave every morning and go off to work while the women stayed around the (freedom) house and cared for the children (read, "students"). Then, "when they [the mostly male voter registration workers] came home you were to be out of the kitchen; . . . they were tired and they had driven long distances and worked under greater pressure."[18]

But it was not simply that the female volunteers did different jobs than the males, but that the jobs typically assigned to them were seen as

[17] Interview with Miriam Cohen Glickman, June 1, 1985.
[18] Letter written by Pam Parker, Allen Papers (in author's possession).

TABLE 5

PREFERRED AND ACTUAL WORK ASSIGNMENTS FOR MALE AND FEMALE VOLUNTEERS

	FEMALES				MALES			
	Preferred		Actual		Preferred		Actual	
	%	N	%	N	%	N	%	N
Freedom schools	48	68	56	196	37	80	33	151
Voter registration	22	31	9	31	32	69	47	212
Community centers	20	29	22	79	8	18	8	38
Communications	6	9	4	15	9	19	4	17
Research	4	6	6	21	5	11	4	16
White communities....	0	0	3	10	6	12	4	18
Legal...	0	0	. . ª		3	6	. . .ª	
Total.....	100	143	100	352	100	215	100	452

NOTE —The volunteers were asked on the applications to rank the order of their preferences for summer work assignments. Their actual work assignments were reported on a list of all applicants accepted as of May 30, 1964 This list was copied from the original, which is contained in the archives of the Martin Luther King, Jr , Center in Atlanta, Georgia
ª Applicants accepted for legal work were not included in the list from which these percentages were calculated.

less important than those the men did.[19] The major distinction here, of course, was between the freedom school teachers and the voter registration workers. One of the voter registration workers characterized the difference in this way:

> I remember the voter registration workers being different from the freedom school people. The voter registration workers were predominantly male . . . [and] adventurous, and they really wanted to do the nitty, gritty [work]. The freedom school people tended to be women and I think . . . with some exceptions . . . tended to be more protective of their persons. . . . So they'd go to a Freedom School and they'd teach the kids about black history. I think it was a very important part of . . . that summer . . . and yet it wasn't the same kind of, if you want, macho adventurism that I was into.[20]

Not surprisingly, a former freedom school teacher remembers things a bit differently: "There was very much a sense [that] . . . voter registration activity was where it was at. And since we had chosen teaching, we were sort of shoved to the side. . . . You know, here [were the] . . . guys running out . . . being macho men . . . you know, 'We're going to go

[19] The generally lower status accorded the female volunteers' work on the project is consistent with Barry Thorne's (1975) discussion of the way women came to do the "shitwork" in the draft resistance movement.
[20] Author's interview with Stuart Rawlings, August 22, 1984.

out and get our heads busted and we'll come back to here where you nurse us . . . and otherwise service us and send us back out again.'"[21]

How are we to reconcile the female volunteers' overwhelmingly positive assessments of Freedom Summer with these forms of sexism? The answer is really quite simple and segues naturally into a discussion of the project's ongoing impact on the volunteers. The fact is the female volunteers' experiences that summer preceded, by several years, the development and widespread dissemination of a contemporary feminist perspective. Lacking that perspective and the language it would later afford them, the female volunteers were not able to name and therefore could not fully *experience* the discrimination that was objectively taking place. Said one volunteer, "I didn't see it [as sexism] . . . at the time. . . . I don't know what I thought of it exactly at the time, but you know sexism was not something that . . . had been made conscious to me at that time. . . . That only came later . . . in the [women's] movement."[22] This and several similar quotes from other volunteers remind us of the consensual and socially constructed nature of experience. Only when viewed through the filter of a later feminist perspective is the summer experienced as "sexist." This may help explain the apparent paradox noted earlier. As near as I can judge, the female volunteers appear to retain much the same positive emotional heritage from that summer as do their male counterparts. It is only in their retrospective assessments of specific aspects of the project that they differ. But as we will see, these differences in retrospective interpretation appear to be the key to explaining a very interesting contrast in the subsequent biographies of the two groups.

Biographical Impact of the Summer on the Volunteers

Elsewhere (McAdam 1988, 1989) I have shown that participation in the summer project has had a number of significant short- and long-term effects on the subsequent activist biographies of the project volunteers. Here I am interested in using gender to decompose those effects. Does the pattern of gender differences evident in recruitment to and participation in the project extend beyond the summer? Has the political impact of the project been mediated by gender?

Several items on the follow-up questionnaire suggest that the female volunteers attribute a greater *effect* to their participation than do the male volunteers. Subjects were asked to describe their "political stance" "before" and "immediately following" the summer using a 10-point

[21] Author's interview with Linda Davis, April 18, 1985.
[22] Author's interview with Jan Hillegas, March 26, 1985.

scale ranging from "1" for "radical left" to "10" for "radical right." Both groups of volunteers ascribed a radicalizing effect to their involvement in the project. But the mean difference in the pre- and postsummer placements on this scale were significantly greater for the female than the male volunteers.[23]

The female volunteers were also more likely to respond negatively to a questionnaire item asking them whether their estimate of six governmental institutions (i.e., the president, Congress, the Justice Department, etc.) had been "raised" or "lowered" as a result of their experiences during the summer. Assigning a value of -1 to a negative response and $+1$ to a positive response, and then summing the responses, yields an estimate of the project's overall political impact on the subject. Once again, the difference in the mean scores on this "impact scale" vary significantly by gender with the female volunteers responding more critically to mainstream governmental institutions than their male counterparts.[24]

These two items tell us that the female volunteers *perceive* the summer as having had a greater impact on them than do the male project veterans. But what about actual behavioral differences? In the subsequent biographies of the male and female volunteers, can we detect a differential imprint of participation in Freedom Summer?

Political Impact of Participation

The evidence from the follow-up surveys paints a complex picture as regards gender and the short-term political impact of participation. One thing is clear: irrespective of gender, involvement in the summer project is clearly related to higher levels of activism in the period 1964–70. But when we add gender to the equation, something very interesting happens. While the female volunteers clearly attribute greater significance to their

[23]

Pre- and Postsummer Estimates of
Political Stance by Gender

	Pre	Post	Net Change
Males	3.49	2.76	$-.73$
Females	3.70	2.60	-1.10

The f-test showed that the difference in these two measures was significant at the 01 level.

[24] The f-test showed that the difference in the mean scores on this "political impact" scale was .001.

summer experiences than do their male counterparts, the behavioral indicators support an opposite conclusion. That is, participation in Freedom Summer appears to have had a much greater impact on the later activist histories of the male rather than the female volunteers. The separate regressions for males and females reported in table 6 clearly support this account.

It is not surprising that the strongest predictor of sixties activism for both males and females is the number of political affiliations reported for the years 1964–70. Of more interest are the other significant relationships shown in table 6. Participation in the Freedom Summer project has a significant effect on sixties activism for the male, but not the female, volunteers. For the females it is their level of activism *prior* to the summer that bears a significant relationship to their level of movement participation for the rest of the decade.

These findings suggest several interesting conclusions that underscore the central importance of gender in mediating the Freedom Summer experience. First, the suggestion in these data is that only women who were previously committed to activism and a broader activist community were likely to make it to Mississippi. This finding appears to confirm a point stressed earlier: to the extent that women faced greater obstacles to participation than men, they needed to be "overqualified"—that is, more experienced and committed to the movement—to overcome the opposition they encountered. Ironically, however, their greater experience and commitment appear to have made the female volunteers *less* susceptible to the kind of radicalizing experience that many of the males appear to have had in Mississippi. So, as table 6 shows, participation in Freedom Summer is a strong predictor of level of sixties activism for the male, but not the female, volunteers. Facing less opposition, the male volunteers did not have to be as sure of their activist commitments prior to the summer and thus were more susceptible to having those commitments deepened as a result of their experiences in Mississippi.

This contrast raises another interesting question: How are we to reconcile the greater *perceived* effect on the female volunteers? This question alerts us to one of the most interesting yet least studied aspects of social movement participation. I am referring to the reconstruction of biography. As recent work on the topic (see Grele 1985; Passerini 1986) reminds us, the reconstruction of biography is a ubiquitous feature of social life. All of us are inclined, in light of changing life circumstances, to recast our self-reflective biographical accounts in subtle or not-so-subtle ways. But intense activism highlights the possibilities for the reconstruction of biography in a way that only a few other life experiences (e.g., divorce, religious conversion, military service) do. To commit to a movement in the way that many of the volunteers did is to undergo a kind of conver-

TABLE 6

RESULTS OF SEPARATE OLS REGRESSIONS FOR MALES AND FEMALES ASSESSING THE
EFFECT OF VARIOUS INDEPENDENT VARIABLES ON THE SUBJECT'S LEVEL OF
ACTIVISM, 1964–70

| | DEPENDENT VARIABLE = ACTIVISM, 1964–70[a] | | | |
| | Males (N = 180) | | Females (N = 105) | |
INDEPENDENT VARIABLE	b	SE(b)	b	SE(b)
Political orientation prior to the summer .	090	.076	063	085
Project participation (yes/no)	1.611*	854	638	2 023
No. of political affiliations, presummer	.316	.349	.633	.676
No. of political affiliations, 1964–70	3.536**	.351	3.340**	440
Level of activism prior to the summer	.203	120	.271*	.126
Constant...	6.801**	1.248	7 316**	1.801

[a] For details of how the dependent variables were coded, see McAdam (1988, p 304)
* $P < 05$
** $P < .01$

sion experience (cf. della Porta, in press). For many, Freedom Summer
came to be the event around which they reconstructed their biographies
in "before" and "after" fashion. This was no less true for the male than
the female volunteers. But the development and application of a feminist
perspective on Freedom Summer in the years following the project has
helped sustain the perceived importance of the project for women in a
way that nothing has for the male volunteers. Let me explore this dy-
namic in a bit more detail. At the close of the summer, the majority of
volunteers—female no less than male—viewed themselves as "move-
ment people." First and foremost this meant the civil rights movement,
though many also were clearly attuned and sympathetic to the emerging
student and antiwar movements. Indeed, they viewed these as one and
the same struggle (McAdam 1988, pp. 162–78). Freedom Summer, then,
loomed large in the volunteers' accounts of how they became "movement
people." But the expulsion of whites from the civil rights movement and
the gradual dissolution of the radical left in the early 1970s slowly eroded
the salience of the designation "movement person." As that identity
became more tenuous, so too did the importance of Freedom Summer as
the pivotal event in the volunteers' reconstructed biographies.

But this dilemma was much more acute for the male volunteers. The
rise of the women's liberation movement served to provide most of the
female volunteers with a highly salient new identity—that of feminist—
around which their biographies could once again be recast. Freedom

Summer was incorporated as a key event in the feminist account of the movement's origins, thus enhancing its significance as a biographical event for the female volunteers.

The feminist account of Freedom Summer runs as follows: the contrast between the egalitarian ideology of the civil rights movement and the forms of gender discrimination evident on the project helped shape an emerging awareness of sexism and the need to organize against it. This view was advanced as early as the fall of 1964 by Mary King and Casey Hayden, two longtime civil rights activists and Freedom Summer participants, in a paper prepared for a SNCC retreat. In this paper King and Hayden drew on events of the summer (and earlier times) to point out and critique women's "second-class" status within SNCC and the movement. This view was subsequently adopted by later feminists and eventually by women's historians such as Sara Evans (1980) and Mary Aiken Rothschild (1982, 1979). Thus, by the time I gathered my data (1984–86), this account of Freedom Summer was firmly established within the feminist community.

But it is not enough that this account was established. The female volunteers would have had to have been exposed to it as well. This raises the question of whether the female volunteers were involved in the women's liberation movement. The answer is clearly yes. Eighty percent of the volunteers reported some level of involvement in the "women's liberation movement" with better than 60% defining their involvement as "very high" or "moderate." So the account linking Freedom Summer with the women's liberation movement was not just established but clearly "available" to the female volunteers as a frame of reference for assessing the impact of the project on their subsequent political biographies. It is hardly surprising, then, that they attribute more significance to Freedom Summer than their male counterparts. But it is the male volunteers whose activist histories during the sixties bear the greater statistical influence of participation in the summer project.

Does this mean that the project had no substantive significance as regards the later political biographies of the female volunteers? Hardly. Clearly their experiences in Mississippi and the ties they forged to one another made the female volunteers ideologically receptive and structurally available to early feminist organizing efforts. What these data *do* suggest is that their *levels* of sixties activism, as opposed to the *forms* this activism took, are not attributable to their involvement in Freedom Summer. Instead, that involvement *as well as* their later sixties activities owes in large part to the same source: the female volunteers' often considerable histories of activism preceding the summer.

But what about long-term effects? If we examine the contemporary political activities of the project applicants, do we still find evidence of

TABLE 7

LEVELS OF CURRENT POLITICAL ACTIVITY BY PROJECT STATUS AND GENDER

	FEMALES		MALES	
	Volunteers	No-Shows	Volunteers	No-Shows
"Very involved" in antinuclear movement (%)	17	12	8	2*
"Very involved" in nuclear-freeze campaign (%)...	11	8	10	3+
Active in at least one social movement (%)	56	42	42	32*
Mean no of political organizations affiliated with	2.5	1.9	1.9	1 1**
Mean score on contemporary activism scale	11.6	9.5	8.5	5.8*

+ P < .10.
* P < .05
** P < .01

the mediating effects of gender? Table 7 suggests a complicated answer to this question.

According to table 7, the female volunteers have remained active relative to the other three groups of applicants. But the male volunteers have since lagged behind the female no-shows. This suggests a diminution of the effect of Freedom Summer participation on male activism that we saw for the earlier period. Meanwhile, the independent effect of gender seemingly has been enhanced. These conclusions are further reinforced by the results of the analysis reported in table 8. When the effects on current activism of gender, project status, and the interaction of the two are simultaneously assessed, only gender remains significant.

Does this mean that the political effects of Freedom Summer participation were confined to the late sixties? The data suggest otherwise. Both the male and female volunteers remain more active than their same-sex counterparts who did not make it to Mississippi. If the level of activism of the male volunteers has declined relative to the late 1960s, the decline is likely the result of the general collapse of the New Left (and the specific vehicle that largely sustained male activism, the antiwar movement) rather than of the withering away of the effects of Freedom Summer. In a similar fashion, the relative increase in activism among the female no-shows should be laid to the general consciousness-raising and mobilizing functions of the women's movement rather than to any diminution of the impact of the Freedom Summer among the female volunteers. The

TABLE 8

OLS REGRESSION PREDICTING CURRENT LEVEL OF ACTIVISM AMONG THE FREEDOM
SUMMER APPLICANTS ($N = 272$)

	DEPENDENT VARIABLE = CURRENT ACTIVISM	
INDEPENDENT VARIABLE	b	SE(b)
No of political affiliations prior to Freedom Summer..............	$-$ 390	.278
Level of activism, presummer...... .	.030	.073
Level of sixties activism..	188**	047
No. of current political affiliations ..	1 186**	.241
Current political orientation . . .	-1.046**	.310
Gender (0 = male, 1 = female) :	2.148*	1 167
Project status (0 = no-show; 1 = volunteer)..	-1 105	1.153
Gender $\times$ status	2.277	1.802
Constant	6.697**	1.746

NOTE —R^2 = 31. For a detailed description of coding the variables, see McAdam (1988)
* $P < $ 05.
** $P < .01.$

gap in the activism levels of the volunteers and no-shows remains intact,
even if other factors have somewhat muted or obscured their effects.

CONCLUSION

The results summarized here underscore the important conceptual point
with which I began this paper. While most work by sociologists still tends
to treat social phenomena as gender neutral, this study again confirms the
powerful way in which gender mediates social life. Quite clearly, Free-
dom Summer was experienced, and continues to be experienced, very
differently by the male and female volunteers. If we are ever to under-
stand human experience and social life, we must, in both research and
theorizing, be attuned to these differences.

This point needs to be made all the more forcefully to those studying
the dynamics of individual activism. While other subfields within sociol-
ogy—sociology of emotions or marriage and the family, for exam-
ple—have made some progress in taking gender seriously as a force in
social life, we have almost totally ignored gender's impact. In doing so
we have perpetuated a fiction: that recruitment to, participation in, and
the consequences of activism are somehow experienced the same by all
participants. Clearly this is false. The research reported here demon-

strates as much. In the case of Freedom Summer, gender mediated every aspect of the experience.

As regards recruitment, the anlysis clearly suggests that the female volunteers faced much greater opposition to their participation than did their male counterparts. Two factors helped those women who made it to Mississippi overcome this opposition. One was organizational sponsorship. The strongest predictor of female participation in the project was the number of organizations the applicants belonged to on the eve of the summer. The other "defense" against the greater obstacles to female participation was a well-established history of activism and involvement in the movement before Freedom Summer. Thus, the female volunteers exhibited one of the qualities traditionally associated with those who break down gender (or racial) barriers: they were "overqualified" for participation in the project.

Gender also clearly differentiated the actual experience of Freedom Summer for the male and female volunteers. Reflecting both strategic considerations and the intersection of traditional work and gender roles, the female volunteers were largely confined to teaching and clerical work on the project, while the men engaged in voter registration and other political activities. The sexual conventions of the era were also evident in Mississippi with the male volunteers and staff members granted far more behavioral license and moral tolerance than their female counterparts. Notwithstanding these forms of discrimination, the female volunteers retain a very positive view of the summer project. Their summer experiences clearly deepened the female volunteers' commitment to the movement while simultaneously increasing their own confidence and self-esteem. The project also exposed them to new political perspectives and forms of community that would, in part through their own efforts, find later expression in the burgeoning student, antiwar, and women's liberation movements.

Gender is also implicated in the subsequent political biographies of the volunteers. Not surprisingly, irrespective of gender, the levels of sixties activism are significantly higher for the volunteers than the no-shows. But only the level of male activism is related to participation in Freedom Summer. For the female volunteers, it is the level of activism prior to and following the summer that is significantly related.

The image is that of a group of women who, on the eve of Freedom Summer, had already transcended traditional gender expectations to a considerable degree. It was this transcedence and their deep commitment to the civil rights movement that enabled them to overcome the greater opposition they faced in applying to the project. More significantly, it is this transcendence as measured by the number of their political affilia-

tions prior to the summer that bears the strongest relationship to the political course of their lives since Freedom Summer.

The male volunteers look quite different. Facing less opposition to their participation, a larger percentage of the males who applied to the project actually made it to Mississippi. Those who did participate had less extensive histories of prior activism and fewer political affiliations than the female volunteers. It is perhaps not surprising, then, that the project appears to have had a stronger behavioral impact on the males than the females. Their level of sixties activism bears the strong statistical imprint of Freedom Summer.

Given the contrast, it is especially interesting to note that the female volunteers *attribute* more personal significance to Freedom Summer than do their male counterparts. This asymmetry in behavior and perception raises the fascinating topic of the reconstruction of biography. Much work remains to be done on the impact of movement participation on the dynamics of biographical reconstruction. For my part, all I can do is offer the following "gendered" account of the process as regards Freedom Summer. The dissolution of the New Left in the early seventies diminished the salience of the identity of "movement person" for the male volunteers, and with it the biographical significance of Freedom Summer in shaping that identity. In contrast, the rise of the women's movement only heightened the significance of Freedom Summer for those female volunteers who identified themselves as feminists. This effect owed to the central role, both positive and negative, assigned to the summer project in the feminist account of the movement's origins. Accordingly, the project has remained more salient in the contemporary biographical accounts offered by the female than the male volunteers.

REFERENCES

Aberle, David. 1966. *The Peyote Religion among the Navajo.* Chicago: Aldine.
Aptheker, Bettina. 1982. *Women's Legacy: Essays in Race, Sex and Class in American History.* Amherst: University of Massachusetts Press
Block, Jeanne H. 1972. "Generational Continuity and Discontinuity in the Understanding of Societal Rejection." *Journal of Personality and Social Psychology* 22:333–45.
Blumberg, Rhoda Goldstein, and Wendell James Roye, eds. 1979. *Interracial Bonds.* Bayside, N.Y : General Hall.
Blumberg, Rhoda Lois. 1980. "White Mothers in the American Civil Rights Movement." Pp. 33–50 in *Research in the Interweave of Social Roles: Women and Men,* vol. 1. Edited by Helena Z. Lopata. Greenwich, Conn.: JAI.
Bookman, Ann, and Sandra Morgen, eds. 1988. *Women and the Politics of the Empowerment.* Philadelphia: Temple University Press.
Brady, David W., and Kent L. Tedin 1976. "Ladies in Pink: Religion and Ideology in the Anti-ERA Movement." *Social Science Quarterly* 56:564–75.

Briet, Martien, Bert Klandermans, and Frederike Kroon. 1984. "How Women Become Involved in the Feminist Movement." Paper presented at the annual meeting of the Society for the Study of Social Problems, San Antonio, Texas.

Cable, Sherry. 1992. "Women's Social Movement Involvement: The Role of Structural Availability in Recruitment and Participation Processes." *Sociological Quarterly* 33.

Canterow, Ellen, ed. 1980. *Moving the Mountain: Women Working for Social Change*. Old Westbury, N.Y.. Feminist Press.

Carden, Maren Lockwood. 1974. *The New Feminist Movement*. New York: Russell Sage.

Chafetz, Janet Saltzman, and Anthony Gary Dworkin. 1986. *Female Revolt*. Totowa, N.J.: Rowman & Allanheld.

Connell, R. W. 1987. *Gender and Power: Society, the Person and Sexual Politics*. Cambridge Polity, in association with Blackwell.

Davies, James C 1969. "The J-Curve of Rising and Declining Satisfaction as a Cause of Some Great Revolutions and a Contained Rebellion." In *Violence in America: Historical and Comparative Perspectives*. Washington, D.C.: Government Printing Office.

della Porta, Donatella. In press "Political Socialization in Left-Wing Underground Organizations: Biographies of Italian and German Militants." In *Participation in Underground Organizations*, edited by Donatella della Porta. Greenwich, Conn.. JAI.

Dill, Bonnie. 1981. "On the Hem of Life: Race, Class and the Prospects for an All-inclusive Sisterhood " Pp. 173–88 in *Class, Race, and Sex: The Dynamics of Control*, edited by Amy Swerdlow and Hannah Lessinger. Boston: G. K. Hall.

Dworkin, Andrea. 1983. *Right-Wing Women*. New York: Perigee.

Dye, Nancy Schrom. 1980. *As Equals and as Sisters: Feminism, the Labor Movement, and the Women's Trade Union League of New York*. Columbia. University of Missouri Press.

Evans, Sara. 1980. *Personal Politics*. New York: Vintage.

Feieraband, Ivo, Rosalind Feieraband, and Betty Nesvold. 1969. "Social Change and Political Violence: Cross National Patterns." Pp. 497–595 in *Violence in America: Historical and Comparative Perspectives*, edited by Hugh David Graham and Ted Robert Gurr. Washington, D.C.: Government Printing Office.

Fendrich, James M. 1977. "Keeping the Faith or Pursuing the Good Life." *American Sociological Review* 42:144–57.

Fernandez, Roberto, and Doug McAdam. 1988. "Multiorganizational Fields and Recruitment to Social Movements." In *Organizing for Social Change: Social Movement Organizations across Cultures*, edited by Bert Klandermans. Greenwich, Conn.: JAI.

Flacks, Richard. 1967. "The Liberated Generation. An Exploration of the Roots of Student Protest." *Journal of Social Issues* 23:52–75.

Foner, Philip. 1979. *Women and the American Labor Movement: From Colonial Times to the Eve of World War I* New York: Free Press.

Fraser, Ronald. 1988. *1968: A Student Generation in Revolt*. New York. Pantheon.

Freeman, Jo. 1973. "Origins of the Women's Liberation Movement." *American Journal of Sociology* 78:792–811.

————. 1975. *The Politics of Women's Liberation: A Case Study of an Emerging Social Movement and Its Relation to the Policy Process*. New York: David McKay.

Friedan, Betty. 1963 *The Feminine Mystique*. New York. W. W. Norton.

Gerlach, Luther P., and Virginia H. Hine. 1970. *People, Power, and Change: Movements of Social Transformation* Indianapolis, Indiana. Bobbs-Merrill

Gerson, Kathleen. 1985. *Hard Choices: How Women Decide about Work, Career and Motherhood*. Berkeley and Los Angeles. University of California Press.

Geschwender, James 1968. "Explorations in the Theory of Social Movements and Revolution." *Social Forces* 47:127–35.

Gilligan, Carol. 1977 "In a Different Voice: Women's Conception of the Self and Morality." *Harvard Educational Review* 47 (4). 481–517.

———. 1982. *In a Different Voice: Psychological Theory and Women's Development.* Cambridge, Mass.: Harvard University Press.

Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78. 1360–80.

Grele, Ronald J 1985. *Envelopes of Sound: The Art of Oral History,* 2d ed. Chicago: Precedent.

Gurr, Ted. 1970. *Why Men Rebel* Princeton, N.J.: Princeton University Press.

Heirich, Max 1977. "Changes of Heart: A Test of Some Widely Held Theories of Religious Conversion." *American Journal of Sociology* 83:653–80.

Hertz, Susan. 1977. "The Politics of the Welfare Mother's Movement: A Case Study." *Signs* 2:601–11.

Hewlett, Sylvia. 1986. *A Lesser Life: The Myth of Women's Liberation in America.* New York: William Morrow.

Hochschild, Arlie. 1983. *The Managed Heart.* Berkeley and Los Angeles: University of California Press.

Hoffer, Eric. 1951. *The True Believer: Thoughts on the Nature of Mass Movements.* New York: Mentor Books, New American Library.

Hole, Judith, and Ellen Levine. 1971. *Rebirth of Feminism.* New York. Quadrangle Books.

Keniston, Kenneth. 1968. *Young Radicals.* New York: Harcourt, Brace & World.

Kessler-Harris, Alice. 1985. "Problems of Coalition-building: Women and Trade Unions in the 1920s." In *Women, Work and Protest: A Century of U.S Women's Labor History,* edited by Ruth Milkman. Boston: Routledge & Kegan Paul.

Klatch, Rebecca 1987. *Women of the Right.* Philadelphia: Temple University Press.

——— 1988. "Coalition and Conflict among Women of the New Right." *Signs* 13:671–94

Kornhauser, William. 1959. *The Politics of Mass Society* Glencoe, Ill.: Free Press.

Lawson, Ronald, and Stephen E. Barton 1980 "Sex Roles in Social Movements: A Case Study of the Tenant Movement in New York City." *Signs* 21:230–47.

Lipset, Seymour M , and Earl Rabb. 1973. *The Politics of Unreason: Right-Wing Extremism in America, 1790–1970.* New York: Harper & Row.

Luker, Kristin. 1984 *Abortion and the Politics of Motherhood.* Berkeley and Los Angeles. University of California Press.

Luttrell, Wendy. 1988. "The Edison School Struggle The Reshaping of Working-Class Education and Women's Consciousness." Pp. 136–56 in *Women and the Politics of Empowerment,* edited by Ann Bookman and Sandra Morgen. Philadelphia: Temple University Press.

McAdam, Doug. 1986. "Recruitment to High-Risk Activism: The Case of Freedom Summer." *American Journal of Sociology* 92:64–90.

———. 1988. *Freedom Summer.* New York: Oxford University Press.

——— 1989. "The Biographical Consequences of Activism." *American Sociological Review* 54:744–60.

McCarthy, John D., and Mayer D. Zald 1973. *The Trend of Social Movements in America: Professionalization and Resource Mobilization.* Morristown, N.J. General Learning.

McCormack, Thelma 1975. "Toward a Nonsexist Perspective on Social and Political Change." Pp. 1–33 in *Another Voice,* edited by Marcia Millman and Rosabeth Moss Kanter. Garden City, N.Y : Anchor.

McCourt, Kathleen. 1977. *Working Class Women and Grass Roots Politics.* Bloomington: Indiana University Press.

Milkman, Ruth. 1980. "Organizing the Sexual Division of Labor: Historical Perspectives on 'Women's Work' and the American Labor Movement." *Socialist Review* 10:95–150.

————. ed. 1985. *Women, Work and Protest: A Century of U S Women's Labor History*. Boston: Routledge & Kegan Paul.

Millman, Marcia, and Rosabeth Moss Kanter 1975 *Another Voice*. Garden City, N.J.: Anchor.

Morgan, Robin, ed. 1970. *Sisterhood Is Powerful: An Anthology of Writings from the Women's Liberation Movement*. New York: Vintage.

Morrison, Denton. 1973. "Some Notes toward Theory on Relative Deprivation, Social Movements, and Social Change." Pp. 103–116 in *Social Movements: A Reader and Source Book,* edited by R. R. Evans. Chicago: Rand McNally.

Oliver, Pamela. 1984. "If You Don't Do It, Nobody Will. Active and Token Contributors to Local Collective Action." *American Sociological Review* 49:601–10.

Oliver, Pamela, Gerald Marwell, and Ruy Teixeira. 1985 "A Theory of the Critical Mass I: Interdependence, Group Heterogeneity, and the Production of Collective Goods." *American Journal of Sociology* 91:522–56.

O'Neill, William. 1969 *Everyone Was Brave: The Rise and Fall of Feminism in America*. Chicago: Quadrangle.

Orum, Anthony M. 1972. *Black Students in Protest*. Washington, D.C.: American Sociological Association.

Passerini, Luisa. 1986. *Fascism in Popular Memory. The Cultural Experience of the Turin Working Class,* translated by Robert Lumley and Jude Bloomfield. Cambridge: Cambridge University Press.

Raines, Howell. 1983. *My Soul Is Rested*. New York: Penguin.

Rochford, E. Burke. 1985. *Hare Krishna in America*. New Brunswick, N.J.. Rutgers University Press.

Rosenthal, Naomi, Meryl Fingrutd, Michele Ethier, Roberta Karant, and David McDonald. 1985. "Social Movements and Network Analysis. A Case Study of Nineteenth-Century Women's Reform in New York State." *American Journal of Sociology* 90:1022–55.

Rothschild, Mary Aiken. 1979. "White Women Volunteers in the Freedom Summers: Their Life and Work in a Movement for Social Change." *Feminist Studies* 5:466–95.

————. 1982. *A Case of Black and White: Northern Volunteers and the Southern Freedom Summers, 1964–1965* Westport, Conn.: Greenwood.

Sacks, Karen Brodkin. 1988. "Gender and Grassroots Leadership " Pp. 77–94 in *Women and the Politics of Empowerment,* edited by Ann Bookman and Sandra Morgen. Philadelphia: Temple University Press.

Sapiro, Virginia. 1983. *The Political Integration of Women: Roles, Socialization and Politics*. Chicago: University of Illinois Press.

Scott, Joan Wallach. 1988 *Gender and the Politics of History*. New York. Columbia University Press.

Snow, David A., Louis A. Zurcher, Jr., and Sheldon Ekland-Olson 1980. "Social Networks and Social Movements. A Microstructural Approach to Differential Recruitment." *American Sociological Review* 45 (5): 787–801.

Snow, David A., and E. Burke Rochford, Jr. 1983. "Structural Availability, the Alignment Process and Movement Recruitment." Paper presented at the annual meeting of the American Sociological Association, Detroit, August

Stacey, Judith, and Barrie Thorne 1985. "The Missing Feminist Revolution in Sociology " *Social Problems*.

Thorne, Barrie. 1975. "Women in the Draft Resistance Movement." *Sex Roles* 2:179–95.

Weitzman, Lenore 1985. *The Divorce Revolution: The Unexpected Social and*

Economic Consequences for Women and Children in America. New York: Free Press.

Wertheimer, Barbara Mayer. 1977. *We Were There: The Story of Working Women in America.* New York: Pantheon.

West, Guida. 1981. *The National Welfare Rights Movement: The Social Protest of Poor Women.* New York: Praeger.

Wiltfang, Gregory L., and Doug McAdam. 1991. "The Costs and Risks of Social Activism: A Study of Sanctuary Movement Activism." *Social Problems* 69:987–1010.

I Will Follow Him: Family Ties, Gender-Role Beliefs, and Reluctance to Relocate for a Better Job[1]

William T. Bielby and Denise D. Bielby
University of California, Santa Barbara

This article tests competing explanations for why wives in dual-earner couples are less willing than husbands to relocate for a better job. Hypotheses are derived from a neoclassical market model that assumes spouses maximize family well-being and a sociological alternative that emphasizes the mediating function of gender-role ideology in decision making by couples. Results from a maximum-likelihood probit model that uses data from the 1977 Quality of Employment Survey indicate that couples' orientation to the "provider role" shapes how they respond to job opportunities. A husband's potential loss from a move appears to deter wives from capitalizing on opportunities at a new location, but a wife's potential loss does not deter husbands. However, differences by gender are substantially smaller among men and women who reject traditional notions about husbands' and wives' roles within families. Implications for future theoretical development and empirical research are discussed.

Geographic relocation among dual-earner couples more often than not results in increased earnings for the husband and decreased earnings for the wife (Mincer 1978). Married women are much less likely than married men to report being willing to move for improved job opportunities (Markham et al. 1983; Markham and Pleck 1986). An employer who knows that familial obligations tie a married employee to her current location need not match higher wage opportunities available in distant locations (Sandell 1977). Thus, it may be that a substantial portion of

the gender gap in wages is due to the dynamics of decision making within families regarding how to respond to opportunities elsewhere.

In this article, we seek to understand why wives are less willing than husbands to relocate for improved job prospects. Specifically, we examine the circumstances under which family considerations lead employed wives and husbands to forgo a move that would lead to job advancement. Of particular interest is whether variation in reluctance to relocate is consistent with rational decision making regarding family well-being. We assess the adequacy of the explanation offered by neoclassical economists regarding family migration decisions and examine whether gender-role ideology mediates husbands' and wives' responses to a job opportunity that requires geographic relocation.

A NEOCLASSICAL MARKET MODEL OF FAMILY MIGRATION

Jacob Mincer (1978), in an article published in the *Journal of Political Economy,* offers a parsimonious and persuasive model of family migration. The central idea is that spouses maximize family well-being and, in doing so, may forgo opportunities that are optimal from a personal calculation of utility maximization. Let G_h and G_w be the potential net gains in utility associated with a relocation opportunity, where the net gain to the family, G_f, is the unweighted sum of G_h and G_w. If both G_h and G_w are positive, the move is optimal for the family as well as for the husband and wife individually. If, however, the husband's net gain exceeds the wife's net loss (i.e., $G_f > 0$, $G_h > 0$, $G_w < 0$, and $|G_w| < G_h$), then the optimal decision for the family is also optimal for the husband, but not for the wife. Thus, if the utility being maximized were family income, the family would relocate if the husband's gain in earnings in the new location exceeded the absolute value of the wife's loss (net of the cost of the move). The wife in this circumstance is, according to Mincer's definition, a "tied mover," since her move is tied to family circumstances that run counter to her "private" calculus. Of course, according to the formal properties of the model, either spouse could be the tied mover.

Conversely, if the wife is faced with a net gain from a relocation opportunity, but the husband's net loss is of greater magnitude (i.e., $G_f < 0$, $G_w > 0$, $G_h < 0$, and $|G_h| > G_w$) then forgoing the relocation maximizes family utility as well as the husband's utility. In this situation, the wife is a "tied stayer"—capitalizing upon her personal utility would make the family worse off, and she stays for the sake of the family. Again, according to the formal properties of the model, either spouse could be a tied stayer.

Mincer extends the model by introducing the possibility of opportunities at more than one alternative geographic location. The location that

maximizes the wife's gain (G_w^{max}) need not be the same as the one that maximizes the husband's gain (G_h^{max}). Yet a third location could maximize family gain $(G_f^{\text{max}} = G_w' + G_f')$, and the move there would lead to forgone private opportunities for both spouses. Both would be tied movers, sacrificing private gains of $(G_h^{\text{max}} - G_h')$ and $(G_w^{\text{max}} - G_w')$, respectively, for husband and wife. In short, both spouses compromise for the sake of maximizing family utility, although one spouse is likely to compromise more than the other. In the language of neoclassical economics, these sacrifices represent "negative private externalities" that are internalized within the family. That is, the discrepancy between the individual's private maximum $(G_w^{\text{max}} \text{ or } G_h^{\text{max}})$ and the gain corresponding to the family optimum $(G_w' \text{ or } G_h')$ measures a cost to the individual that is recompensed within the family.[2]

The neoclassical model is symmetric with respect to spouses. In terms of its formal properties, husbands and wives are treated identically. The neoclassical model does *not* assume that the husband's potential gains from moving are more important than the wife's in decisions regarding migration (Sandell 1977). Instead, the model assumes that each spouse's potential gain or loss is weighted equally in the computation of family well-being. Nor does the neoclassical model assume that the husband is able to impose his own private interests on the family. Instead, the model assumes that each spouse places family well-being ahead of personal well-being. According to the model, private losses fall disproportionately on wives solely because of the structure of the labor market. That is, the labor market is structured such that husbands' gains from opportunities elsewhere tend to exceed wives' losses from moving, so tied movers are disproportionately female. And, wives' gains from opportunities elsewhere seldom exceed husbands' losses from moving, so tied stayers are also disproportionately female.

Considerable empirical evidence has been produced that is consistent with predictions from Mincer's model. For example, dual-earner families migrate less than those with a single earner (Mincer 1978). Further, on average, husbands' earnings increase after migration, while wives' earnings and employment levels decline, at least initially, consistent with wives' disproportionate representation among tied movers (Long 1974; Da Vanzo 1976; Duncan and Perrucci 1976; Mincer 1978; Lichter 1980; Spitze 1984; Maxwell 1988).

---

[2] In a further elaboration of the model, following Becker (1974), Mincer introduces the net gains from marriage for each spouse, M_h and M_w. He then shows that the marriage dissolves (i.e., one spouse relocates and the other does not) when migration ties, $(G_h^{\text{max}} - G_h') + (G_w^{\text{max}} - G_w')$, exceed the net gains from marriage, $M_h + M_w$. In this situation, private externalities are not internalized within marriage.

SOCIOLOGICAL APPROACHES: RELATIVE RESOURCES AND COUPLES' DECISION MAKING

Social exchange theory (Emerson 1976) applied to husbands' and wives' decision making (Blood and Wolf 1960; Scanzoni 1970; McDonald 1980; Blumstein and Schwartz 1983; England 1989) yields predictions similar to those of Mincer's neoclassical model. If financial resources provide leverage in bargaining between spouses, then the partner with greater earnings capacity is likely to gain the most in negotiating over whether or not to relocate for a job opportunity in a different location.[3] Thus, both the neoclassical and social exchange theory perspectives predict that married women are more likely than married men to be tied movers and tied stayers because of women's lower earning potential. Also, both perspectives are symmetric with respect to marital partners. That is, both predict that husbands are more likely than wives to be tied movers and tied stayers whenever a wife's earnings capacity exceeds her husband's. Since the two approaches have similar empirical implications, we will not attempt to distinguish between neoclassical and social exchange explanations in our analyses.

However, neoclassical and social exchange explanations do differ in one respect. Mincer's neoclassical model is based on utility maximization. That is, an individual forgoes a personal gain because he or she derives greater utility from enhancing family well-being than from personal well-being. In contrast, social exchange theory invokes the notion of *power* as the mechanism through which decisions are made. That is, the spouse in command of the most resources is able to impose outcomes that further her or his own goals to the detriment of the partner's (Hood 1983, pp. 17, 178).

In recent years, the social exchange or "relative resource" theory (Blood and Wolf 1960) approach to husbands' and wives' decision making and marital power has been criticized and supplemented by approaches that emphasize the *process* of decision making by couples and the mediating role of gender-role ideology in that process (Young and Wilmott 1973; Hood 1983; Szinovacz 1987; Godwin and Scanzoni 1989a). Below, we show how this work suggests an alternative explanation of the tied stayer phenomenon that challenges the symmetry of both neoclassical and social exchange theory approaches.

---

[3] In an elaboration of this perspective, England and Farkas (1986) note that homemakers bring nonfinancial resources to a marriage, but these resources are largely "relationship specific." Whereas financial resources from employment can be transferred to a new marital relationship at little or no cost, the homemaker's relationship-specific resources are of little value in a new marriage. Thus. the partner in the earner role has more alternatives than the homemaker, and therefore the homemaker's nonfinancial resources contribute little to that partner's power within the marriage.

SOCIOLOGICAL APPROACHES: GENDER-ROLE IDEOLOGY AND THE PROVIDER ROLE

Neoclassical and social exchange approaches ignore the household *roles* husbands and wives occupy, the gender-role *beliefs* they subscribe to regarding those roles, and the effect of these beliefs on both the process and outcome of couples' decision making. Hood (1983, p. 5) defines *family roles* as "mutual expectations negotiated by the actors that define each actor's responsibility to other family members in a given context." Neoclassical and social exchange models implicitly assume that such expectations are perfectly flexible and adapt seamlessly to changes in the opportunities and costs facing household members. In fact, even in contemporary American families and despite wives' recent gains in the paid labor force, husbands more often than not assume the provider role (Bernard 1981; Haas 1986; Thompson and Waller 1989; Perry-Jenkins and Crouter 1990). Further, whether or not a spouse assumes (or shares) responsibility for the provider role may shape the value placed on that spouse's earnings potential in bargaining over the division of family roles and responsibilities. For example, in a detailed qualitative study of 16 dual-earner couples, Hood (1983) found that a husband who views his wife as a coprovider is more willing to take on shared responsibility for parenting and household work than is a husband who sees himself as primarily responsible for the provider role. It may also be true that an employed wife who is not recognized as a coprovider cannot rely on her earnings potential as a resource in bargaining over paid work activities.

In short, gender-role ideology introduces asymmetry into the process by which husbands and wives decide how to respond to a job opportunity in a different location. As Hood (1983, p. 7) hypothesizes, a spouse's bargaining power is shaped by "the mutually recognized right or authority to exercise power in a given area" as well as by relative resources. Thus, when the provider role is defined as the husband's responsibility, the wife's net economic gain (or loss) from a prospective geographic move is likely to be discounted relative to that of the husband. If so, then an employed wife is more likely to be either a tied mover or a tied stayer than is an employed husband, regardless of how much parity or imbalance in the economic resources each spouse brings to the relationship.

Asymmetry in the job relocation process is suggested in empirical research by sociologists. Duncan and Perrucci (1976), analyzing longitudinal data on college graduates, found that wives' occupational prestige and proportionate contributions to family income—factors that should contribute to tied-stayer status among husbands—did *not* deter family migration. Lichter (1983), analyzing longitudinal data on a national sample of women, found that wives' returns to migration did *not* increase with education and occupational status (measured both absolutely and

relative to husbands' schooling and status). Thus, these findings at least hint at the possibility that the wife's economic interests are discounted in relocation decisions, even when acting upon those interests would increase the overall economic well-being of the family.

However, none of the existing empirical research was designed explicitly to test the neoclassical model of family migration against sociological alternatives. Moreover, most of the existing research on gender ideology, the provider role, and household decision making focuses on the division of labor and responsibility regarding household roles (e.g., Hood 1983; Hardesty and Bokemeier 1989) rather than on work activities and geographic mobility. The research presented below is a first step toward a direct evaluation of neoclassical versus sociological approaches to the issue of spouses' responses to job opportunities that require geographic relocation.

RELUCTANCE TO RELOCATE: HYPOTHESES DERIVED FROM THE NEOCLASSICAL PERSPECTIVE AND SOCIOLOGICAL ALTERNATIVES

Our focus is on the factors that predispose employed husbands and wives to be tied stayers. Specifically, we will examine the circumstances under which individual husbands and wives report a reluctance to relocate for a much better job because of family obligations. Hypotheses derived from the neoclassical perspective on migration are listed in table 1. Job- and firm-specific investments are hypothesized to tie an individual to his or her current location. These include factors such as job-specific skills, job security, and firm-specific benefits that one sacrifices when leaving one job and employer for another. In contrast, general skills—abilities that are as valuable to other employers as they are to an individual's current employer—should reduce one's reluctance to relocate for better

TABLE 1

DETERMINANTS OF RELUCTANCE TO RELOCATE FOR IMPROVED JOB OPPORTUNITIES: HYPOTHESIZED RELATIONSHIPS DERIVED FROM A NEOCLASSICAL MARKET MODEL

Variable	Hypothesized Relationship
Job-specific investments	+
Firm-specific investments	+
General skills	−
Spouse's job investments	+
Location-specific family investments	+
Female main effect	0

NOTE.— + = positive relationship; − = negative relationship; 0 = no relationship

job opportunities. Family traits tying an individual to his or her current location are spouse's job-specific investments and family investments in location-specific resources such as schools, child care, and the like. A spouse's forgone earnings, the expense of his or her job search in a new location, and the monetary and psychic toll of relocating children to new schools, neighborhoods, and child-care facilities are additional costs associated with a move.

Overall, women are expected to report a greater reluctance to relocate for a better job because of family obligations. As noted above, according to the neoclassical model, women are more likely than men to be tied stayers because wives' potential gains from a move seldom exceed their husbands' losses. However, if the factors listed in table 1 adequately capture the costs involved in a relocation for a better job, then there should be no *net* difference between men and women in the propensity to report a reluctance to relocate because of family obligations. That is, compared with working husbands faced with comparable benefits and costs from a move, employed wives should be neither more nor less likely to express a reluctance to relocate because of family considerations. Of course, any small net difference could easily be attributable to model misspecification or imperfect measures. However, a modest to strong *net* effect in the direction of women being less willing to relocate would be difficult to reconcile with a model that assumes that husbands' and wives' gains and losses are weighed equally in decisions about job moves.

A sociological approach that introduces the mediating function of gender-role ideology suggests supplementing the neoclassical model with the asymmetries hypothesized in table 2. If husbands discount the value of their wives' paid work to family well-being, then spouses' job investments should have little or no effect on deterring a willingness to relocate for a better job among employed married men. That is, the predictions of the neoclassical model regarding spouses' job investments should apply only to employed married women.

TABLE 2

DETERMINANTS OF RELUCTANCE TO RELOCATE FOR IMPROVED JOB OPPORTUNITIES:
HYPOTHESIZED ASYMMETRIC RELATIONSHIPS

	HYPOTHESIZED RELATIONSHIP	
VARIABLE	Males	Females
Spouse's job investments	0	+
Nontraditional gender-role ideology	+	−

NOTE.— + = positive relationship, − = negative relationship; 0 = no relationship

Moreover, dual-earner couples differ in the extent to which they subscribe to traditional gender-role ideology regarding the husband's role as provider. Husbands and wives with progressive values regarding work and family roles should be more flexible with regard to sharing responsibility for the economic well-being of the family (Raush 1977). Thus, while a traditional husband might be inclined to pursue job advancement despite the consequences for other family members, a nontraditional husband should be more inclined to express a reluctance to do so because of family obligations. In addition, while an employed wife with traditional values regarding family roles might be reluctant to pursue a job advancement that might threaten her husband's role as provider, a nontraditional wife should be more inclined to pursue job advancement that enhances the economic well-being of the family. Accordingly, we hypothesize that nontraditional gender-role ideology will have a positive effect on husbands' reluctance to relocate for a better job because of family obligations and a negative effect on wives' reluctance to do the same.

Below, we first describe our data and measures. Then, for individuals in dual-earner couples, we estimate a statistical model of reluctance to relocate because of family considerations, and we interpret results with respect to the hypotheses derived above. In the final section we assess the implications of our findings for competing perspectives on gender and couples' decision making and make suggestions for further research.

DATA AND MEASUREMENT

Data

Data are from the 1977 Quality of Employment Survey (QES), a representative multistage probability sample of adults who are at least 16 years old and who live in households and work at least 20 hours per week (Quinn and Staines 1979). An 80-minute personal interview was completed by 1,515 individuals (79% of the eligible respondents). Our analyses are restricted to married respondents in dual-earner families and are based on a sample of 162 women and 197 men with complete data on the variables included in our statistical models. Detailed descriptions of independent variables appear in the Appendix in table A1, and descriptive statistics for all measures are reported in table 3. These data are publicly available from the Inter-University Consortium for Political and Social Research.

Reluctance to Relocate Due to Family Considerations

Willingness to relocate for improved job opportunities was assessed in the QES with the following item: "Suppose you were offered a job that

TABLE 3

DESCRIPTIVE STATISTICS, 1977 QUALITY OF EMPLOYMENT SURVEY, BY SEX FOR
MARRIED RESPONDENTS WITH EMPLOYED SPOUSES

| | MALES | | FEMALES | |
VARIABLE (RANGE)	Mean	SD	Mean	SD
Dependent variable:				
Reluctant to relocate, any reason (0–1)	.57		.89	.
Reluctant to relocate, family reasons (0–1)	16	.	.56	. . .
Job- and firm-specific investments·				
Job security (1–4)	3 13	.76	3 34	.63
Stake in job (1–4)	2.73	.88	2 33	.91
Skill utilization (1–4)	2.83	82	2 78	79
Specific training, in years (0–10)	2.60	2.09	1 54	1 40
Firm size, in 100s of employees (.05–30)	6.38	10.21	3.83	7.22
Self-employed (0–1)	.11	. . .	.09	
Earnings, in $1,000s (1–97)*	16 97	10 26	8.00	4.71
Intrinsic rewards (1–4)*	2 82	.68	2.99	.69
General skills:				
Part-time (0–1)	.05	. .	16	. . .
Part-year (0–1)	.11	. . .	23	. . .
Work discontinuity (0–.97)	11	.16	31	˙˙ 25
Occupation proportion female (0–1)	20	.25	.69	30
Laborer (0–1)	.04	.	.02	. .
Education (6–18)	13.04	2 79	12.72	2.50
Age (18–78)*	40 27	11 68 ˋ	36.94	11 40
Spouse's job investments·				
Spouse's earnings, in $1,000s (0–61)	7.63	8.34	14.32	7.70
Spouse's specific training, SVP (0–10)	1.62	1.53	2.51	1.90
Family investments				
Children in household (0–1)	.61	. .	61	. . .
Responsible for child (0–1)*	.07	. .	49	. . .
Years married (1–55)*	14 68	11.46	14.36	11 21
Gender-role ideology.				
Nontraditional beliefs (0–1)	.25	.	43	. . .

NOTE —Correlation matrices are available from the authors upon request SDs are not reported for binary variables. N = 197 males, 162 females
* Control variable, no directional hypothesis specified.

is *much better* than your present one, but located in another community at least 100 miles away. How willing would you be to move to the other community to take this job?" The possible responses are "not at all willing," "not too willing," "depends on the circumstances," "somewhat willing," and "very willing." A response in any of the first three categories was followed with a probe: "Why wouldn't you be willing to move?" From the open-ended responses, up to three reasons were coded and classified into various categories, including one for "family ties,

family obligations" and another for "employment of others in house-hold."[4] The dependent variable for our analyses is a binary item coded "1" for respondents who express some reluctance to move *and* who list family factors as a reason for that reluctance and "0" otherwise. We assume that the binary responses reflect an underlying unobserved con-tinuum of reluctance to relocate because of familial ties. Accordingly, statistical analyses are based on a binary probit model (Winship and Mare 1984). Coefficients in the probit model reflect the effects of work and family traits on a respondent's propensity to be tied to her or his current location by family considerations, or, in Mincer's (1978) terminol-ogy, the propensity to be a tied stayer.

One advantage of this particular measurement is that both men and women are confronted with the same option regarding a hypothetical job opportunity elsewhere: a "much better" job located 100 miles away. In contrast, in analyzing the economic consequences of actual moves (e.g., Lichter 1983; Maxwell 1988), there is no way of knowing about differ-ences in the spatial distribution of job opportunities facing married men and women. Nor can an analysis of actual moves unambiguously sepa-rate migration motivated by job opportunities from moves motivated by housing and neighborhood considerations or other factors not directly related to one's job.

On the other hand, one limitation of our measure is that men and women may have different standards regarding what constitutes a "much better" job. Ideally, it would have been preferable to have re-sponses to survey items proposing specific absolute (e.g., $5,000) or rela-tive (e.g., a 15% increase) gains. Nevertheless, our models *do* control for attributes of the respondent's current job, labor supply, and human capi-tal. Any sex difference in standards regarding what constitutes a "much better" job *net* of these factors is likely to be small.

Independent Variables

Job- and firm-specific investments.—Among our measures are percep-tions of job security, having a sense of a "stake" in one's job, and having one's skills fully utilized on the job. The scaling of these measures and the survey items upon which they are based are described in table A1.

Specific training is measured by the average "specific vocational prepa-

[4] The QES codebook lists as examples of "family ties or obligations" responses such as "My family is here," "I'm married," "My husband would hate to move," and "It's a bad time for my kids to change schools." Examples of "employment of others in household" are responses such as "I wouldn't want my husband to leave his job" and "My husband makes 10 times what I can make."

ration" (SVP) score of the respondent's three-digit census occupation. As an estimate of "the amount of time required to learn the techniques, acquire the information, and develop the facility needed for average performance in a specific job-worker situation" (Roos and Price 1981, p. 9), it serves as a proxy for a respondent's job-specific training.

Firm-specific benefits and advancement opportunities are greater in large establishments than in small ones (Brown, Hamilton, and Medoff 1990). Accordingly, we include establishment size (number of employees) among our measures of job- and firm-specific investments. Since self-employed individuals are likely to have made site-specific investments in their businesses, we include self-employment status among our measures as well.

Both intrinsic job rewards and earnings are included among our measures of job- and firm-specific investments. These measures are included as control variables, and we offer no hypotheses about their effects on reluctance to relocate because of family considerations. On the one hand, if earnings and the level of intrinsic rewards are truly attributes of the *job* and not of the incumbent, then they are location-specific and should deter relocation. On the other hand, they may in part represent attributes of the incumbent that are marketable elsewhere, thereby facilitating relocation. Accordingly, we offer no hypotheses about these two measures.

General skills.—The skills of those who work part-time are likely to be generalizable to different locations, thereby facilitating relocation. Similarly, individuals who anticipate discontinuous labor-force participation are likely to choose jobs that are easy to exit and reenter, regardless of location. Accordingly, we include part-time and part-year work and discontinuity of labor-force experience as indicators of general skills hypothesized to have a negative effect on reluctance to move (cf. Markham et al. 1983).

According to human capital explanations of occupational sex segregation, female-dominated occupations are those that penalize workers least for turnover and time out of the labor force (Polachek 1976; cf. England 1982, 1984). Skills acquired in these occupations should be easily transferable, and therefore the proportion female in a respondent's three-digit occupation category is included among our measures of general skills.

Compared with other occupations, unskilled manual labor involves few firm-specific skills. In preliminary models, we included binary variables to capture differences among 10 major occupational categories. However, the only consistent differences were between laborers and all other occupations. Accordingly, only the binary variable for laborers is included in the results reported below.

Following Markham et al. (1983), we expect individuals with more schooling to have knowledge about distant opportunities and skills that

are valued elsewhere. Therefore schooling is included among our measures of general skills.

Age, listed along with our measures of general skills, is included in our statistical models to control for life-cycle variation in the propensity to move. We control for linear and quadratic effects of age but offer no specific hypotheses about those effects.

Spouse's job investments.—Following Mincer (1978), we expect spouse's income to be a strong deterrent to relocation. The higher the spouse's income, the less likely it is that the potential gain from a move for the respondent would offset the losses for the spouse. Conversely, the economic loss to the spouse resulting from a move is minimal if the spouse has very low earnings. Ideally, the potential loss facing the spouse would be measured directly. Nevertheless, the amount of income at risk is obviously greater for respondents whose earnings are considerable compared to those whose spouses earn very little.

Finding employment in a new location for a spouse with specific training should be more difficult than doing so for a spouse with more generalized skills. Our measure is the average SVP score for the spouse's three-digit occupational category.[5]

Family investments.—As noted above, costs, both material and emotional, are involved in moving children from one neighborhood to another. New arrangements have to be made for schooling, child care, and recreation, and children must establish new friendship networks. Accordingly, our models include the presence of children in the household as a measure of family investments in one's current location. Our model also included a binary measure indicating whether the respondent has primary responsibility for child care. According to the neoclassical model, husbands and wives are assumed to be equally altruistic toward their children. Consequently, according to that model, whether an individual has primary responsibility for child care should have no net impact on reluctance to relocate, and we hypothesize no relationship. Finally we also include the number of years the respondent has been married as a control variable.

Gender-role ideology.—We classify respondents as holding nontraditional gender-role beliefs if they do not believe that a husband should be sole provider for his family and they do believe that a mother's relationship with her children does not suffer if the mother is employed outside the home. Individuals who are neutral or express contrary views on either

[5] Preliminary models also included spouse's self-employment status, spouse's education, and spouse's major occupational category. However, these traits had no effect on reluctance to relocate because of family considerations, net of spouse's earnings and specific training.

issue are classified as holding traditional gender-role beliefs. The specific survey items upon which our classification is based are described fully in table A1.

RESULTS

Gender Differences in Reluctance to Relocate

A majority of the respondents, 57% of the males and 89% of the females, expressed at least some degree of reluctance to relocate for a better job (table 3). Consistent with each of the theoretical perspectives summarized above, women in dual-earner couples were substantially more likely than men to cite family considerations as a reason for their reservations about relocating for a better job. Over half of the women (56%) reported a reluctance to relocate because of family considerations, compared with just 16% of the men.[6]

Descriptive statistics in table 3 show several gender differences in investments and skills that are consistent with the neoclassical explanation of the tied-stayer phenomenon. In general, husbands in dual-earner couples have higher levels of job- and firm-specific investments than do wives. In addition, women have higher levels of general skills. Specifically, compared with women, men work for larger firms, have more specific training, and report a greater stake in their current jobs. Women are more likely to work part-time or part-year, have discontinuous patterns of labor-force participation, and (of course) work in female-dominated occupations, traits associated with general skills. Spousal earnings, are, on average, nearly twice as high for women as for men. Accordingly, the amount of spouse's earnings at risk from a move is considerably higher for women than for men. While these differences are consistent with the neoclassical model, it remains to be seen whether they account for the substantial difference between men and women in reluctance to relocate. Results from the statistical models discussed below examine whether this is so.

Determinants of Reluctance to Relocate among Men and Women in Dual-Earner Couples: Evaluating Alternative Perspectives

The probit model is estimated in two steps. Model 1 includes the independent variables relevant to the neoclassical explanation of the tied-stayer

---

[6] Besides family considerations, reasons most often cited for being reluctant to relocate were: the attraction of the respondent's current community, inertia (e.g., "We're established here"), and, especially for men, financial ties such as home ownership.

phenomenon: measures of job- and firm-specific investments, general skills, spouse's job investments, and family investments. It also includes a binary variable for gender. If gender differences in the costs and benefits of relocating for a better job are adequately represented by the model, then the coefficient for "female" should be small. A large net effect of female would indicate that wives in dual-earner couples are more likely to forgo a move because of family considerations than are husbands with the same kinds of job investments, skills, and family situations. In other words, such a finding would indicate that gender differences in the tied-stayer phenomenon cannot be explained fully by couples' calculations of the potential net economic gain to the family resulting from a relocation for the job advancement of one spouse but not the other.

Model 2 adds the effect of gender-role ideology and introduces interaction effects that capture the asymmetries hypothesized to follow from the impact of beliefs about the provider role on couples' decision making about relocation for a better job. Strong interaction effects consistent with our hypotheses would indicate that the gendered structure of family roles sharply constrains the process by which husbands and wives evaluate their relative contribution to a family's economic well-being.

A logarithmic (base e) metric is used for respondent's and spouse's earnings. Because of the interactive specification, the effects of gender and nontraditional beliefs are contingent upon spouse's earnings. To facilitate interpretation, the variable log of spousal earnings has been deviated from its sample mean.[7] Accordingly, main effects of gender and nontraditional beliefs represent net differences evaluated at the mean of (logged) spousal earnings.

Consistent with the neoclassical explanation of family migration, estimates for model 1 in table 4 show that both a spouse's earnings and a spouse's specific training contribute to a reluctance to relocate because of family considerations. In addition, skill utilization, an indicator of job-specific investments, has a significant effect in the hypothesized direction.[8] However, these effects are tiny in comparison to the effects of earnings and gender. Compared with those with low earnings, well-paid individuals are much less likely to let family considerations interfere with a potential relocation for a much better job. If we assume that those with high earnings benefit more from such a move in the long run, then this

[7] The mean of log spousal earnings is 8.9636, corresponding to $7,813 in the original (dollar) metric.

[8] Overall, the probit results are roughly consistent with those of Markham and Pleck's (1986) OLS regressions of the QES "willingness to move" measure on several of the same variables included in our specification.

TABLE 4

Determinants of Reluctance to Relocate because of Family Considerations· Maximum-Likelihood Probit Estimates

Independent Variable	Model 1	Model 2
Job- and firm-specific investments.		
Job security.	.14	.13
Stake in job	− 07	− 03
Skill utilization	.19*	.25**
Specific training (SVP).	−.03	−.02
Firm size (in 100s of employees).	.00	00
Self-employed	−.39	− 43
Log earnings	− 52***	−.55***
Intrinsic rewards	−.13	−.14
General skills·		
Part-time	−.10	−.12
Part-year	−.16	−.15
Work discontinuity	−.23	−.34
Occupation proportion female	−.15	−.24
Laborer	−.66	−.50
Education	.03	02
Age (in 10s of years)	55	−.14
(Age − 40)²	−.05	04
Spouse's job investments:		
Log spouse's earnings	16*	−.31
Spouse's specific training (SVP)	.10**	.07
Family investments		
Children in household	13	.20
Responsible for child	− 10	−.16
Years married (10s)	−.05	− 07
Gender-role ideology		
Nontraditional beliefs		68***
Female main effect	74***	1.06***
Interactions:		
Female × log spouse's earnings		1 23***
Female × spouse's specific training		.03
Female × nontraditional		−1 18***
Male × nontraditional beliefs × log spouse's earnings		.54*
Constant.	1.88	3 03
L^2	369.25	339.29

Note —Significance levels are for one-tailed tests except for "log earnings," where a two-tailed test with no directional hypothesis was used

* $P < .10$.
** $P < 05$.
*** $P < 01$

finding is consistent with a neoclassical approach toward maximizing family well-being.[9]

In contrast, the large net gender difference is difficult to explain from a neoclassical perspective. Holding constant job- and firm-specific investments, general skills, spouse's job investments, and family investments accounts for just 40% of the large overall gender difference in reluctance to relocate because of family considerations.[10] In other words, according to model 1, compared with married women in similar work and family situations, married men are substantially less likely to express reservations about familial obligations when contemplating a move for a better job.

The large effect of gender in model 1 coupled with the small effects of every other factor but that of earnings seems to suggest that married men are almost always more likely than married women to pursue personal gain in the sphere of paid employment, regardless of the consequences for family well-being. However, model 2, which acknowledges that gender-role ideology shapes how married men and women respond to job opportunities, shows that determinants of the tied-stayer phenomenon are more subtle and complex.

As noted above, the neoclassical model is implicitly symmetric. It assumes that women's disproportionate representation among tied stayers is due to differences between husbands and wives in the *levels* of cost and gain factors, not due to differences in the *effects* of these factors. Results for model 2 in table 4 show conclusively that the implicit symmetry of the neoclassical model fails to fit the data adequately. Introducing the binary variable for gender-role beliefs and the four variables capturing interaction effects reduces L^2, the likelihood-ratio chi-square statistic, by nearly 30 points. A reduction of that magnitude with five degrees of freedom allows us to reject the hypothesis that the five additional coefficients are jointly zero with a type-I error rate of less than .001. Moreover, the pattern of individual coefficients corresponds almost exactly to our hypotheses: spousal earnings are more consequential for women than for men, and the model's asymmetries are much more pronounced for those

[9] In other words, the dollar value of the future earnings generated by a "much better job" is likely to be considerably higher for a person who moves from an already highly paid position than for a person moving from a poorly paid job.

[10] Evaluated at a probability of .16 (the male mean), the probit coefficient for the female main effect in model 1 corresponds to an effect of .24 in the probability metric. In other words, compared with a man whose traits imply a probability of expressing reluctance to relocate because of family considerations of .16, a woman with identical traits would have a probability of .40 of expressing the same view. Since the zero-order gender difference is .56 − .16 = .40, variables other than gender mediate 100% × (.40 − .24)/.40 = 40% of the total difference.
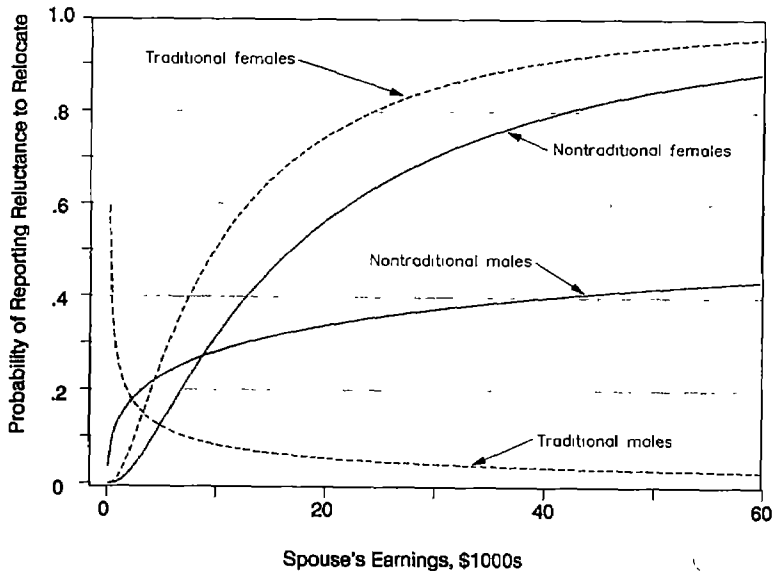
FIG. 1.—Reluctance to relocate due to family considerations, by sex, spouse's earnings, and gender-role beliefs.

with traditional gender-role beliefs than for those with nontraditional beliefs.

The impact of gender-role beliefs and the asymmetries of the tied-stayer phenomenon are apparent in figure 1. The figure shows the probability of reporting a reluctance to relocate because of family considerations as a function of spouse's income, separately for traditional males, traditional females, nontraditional males, and nontraditional females, based on the coefficients for model 2. (Predicted probabilities are computed with all other variables set to their mean values.)

The gender difference in the effect of spouse's earnings is greatest among working husbands and wives who hold traditional gender-role beliefs. Among traditional wives in dual-earner couples, the propensity to be a tied stayer (as indicated by our measure) increases sharply with spouse's earnings, exactly as predicted by the neoclassical model. Their response pattern is consistent with a maximizing approach to family economic well-being. Traditional wives whose husbands earn very little do not report family ties as an obstacle to moving to get a better job. However, the greater their husbands' earnings (and, presumably, the greater the potential for lost earnings caused by a geographic relocation), the more likely these women are to express a reluctance to relocate because of family considerations.

In contrast, the response curve for traditional husbands is essentially flat. (The spike as spouse's earnings approach zero lies outside the range of spouse's earnings, and it is an artifact of the logarithmic specification coupled with a negative but insignificant slope coefficient.) In other words, traditional husbands rarely express any reluctance to move because of family considerations, regardless of how little or how much their spouses earn. The loss of part or all of a wife's contribution to family income is no deterrent to relocation for these husbands, even when the amount of income at risk is relatively large.

Gender differences in the effects of spouse's earnings are less pronounced among husbands and wives holding nontraditional beliefs. As hypothesized, holding such beliefs has a net effect of reducing the propensity for women to report a reluctance to relocate because of family considerations and *increasing* the propensity of men to do the same. Unlike traditional married men, nontraditional husbands are more strongly tied to their current locations when their wives are being relatively well paid.

Nevertheless, gender differences do exist among respondents with nontraditional gender-role beliefs. Nontraditional wives are more sensitive to their husbands' earnings than nontraditional husbands are to their wives' earnings. Moreover, at higher levels of spousal earnings, nontraditional females are more likely to report a reluctance to relocate because of family considerations than are nontraditional males with comparable work and family situations (although the gender difference is substantially less than that among respondents with traditional beliefs).[11]

In sum, our results show that gender difference in the tied-stayer phenomenon among husbands and wives in dual-earner couples is fundamentally shaped by gender-role beliefs. The neoclassical model's explanation of the difference in terms of couples' maximization of family utility fails to capture the asymmetry in how husbands' and wives' contributions are valued. A man with traditional gender-role beliefs appears to give little consideration to the disruption of his wife's employment and her potential loss of earnings when evaluating a potential job opportunity for himself in a new location. Moreover, this orientation among tradi-

---

[11] According to the point estimates for model 2, the response curves in fig. 1 for nontraditional respondents cross at a level of spouse's earnings of $9,300, slightly above the median of spouse's earnings at $8,500. In contrast, the curves for traditional respondents cross at $1,681, which corresponds to the seventh percentile of the spousal earnings distribution. Thus, the model predicts that traditional females will express a greater reluctance to relocate than traditional males throughout almost the entire range of the distribution of spousal income. But among nontraditional respondents, females are actually predicted to be less likely to express a reluctance to relocate due to family considerations across a range of spousal earnings that spans the lower half of the distribution of that trait.

tional males persists even when the level of wife's earnings at risk is substantial.

In contrast, gender differences among those who reject traditional notions of women's and men's roles at home and at work are less pronounced. These men and women, who hold more flexible views toward responsibility for the family's economic well-being, appear to value their own and their spouse's work roles more equally. They are more likely to express an inclination to forgo a move to improve their own job opportunities when it implies that their spouses will have to give up highly paid jobs. But even this propensity is more characteristic of nontraditional females than of nontraditional males. So, as measured here, gender-role beliefs mediate some, but not all, of the asymmetry in the tied-stayer phenomenon.

Our findings have implications for notions of power, ideology, and decision making within families, and they suggest some interesting issues for future research. We address these issues in the concluding section.

CONCLUSIONS

Marital Power, Economic Resources, and Gender-Role Ideology

Our findings suggest that as recently as the late 1970s, a man with traditional gender-role beliefs would refuse to let his spouse's job interfere with his own job advancement. In contrast, a woman with the same beliefs and opportunities would apparently sacrifice job advancement if it meant asking her husband to leave a well-paying job. Our findings for respondents with traditional gender-role beliefs are consistent with a decision-making process wherein men pursue their own self-interest, while women evaluate what is best for the economic well-being of the family. In terms of the neoclassical model, women with traditional beliefs and well-paid spouses anticipate being tied stayers, but traditional men in the same situation do not.

The findings for respondents with traditional beliefs challenge both Mincer's neoclassical model and the social exchange approach to power, relative resources, and decision making within marriage. The neoclassical model of family migration might be salvaged by assuming well-being $(G_f = G_h + G_w)$.[12] However, this modified model no longer provides a neoclassical explanation for gender differences in the tied-stayer phenomenon, since asymmetry has been imposed from outside the model. More-

---

[12] Alternatively, wives with traditional beliefs might weigh husbands' economic gains more heavily than their own in computing family well-being; i e., $G_f = aG_h + bG_w$, where $a > b$.

over, additional ad hoc adjustments would be required to account for the behavior of men and women with nontraditional gender-role beliefs.

Similarly, the social exchange approach cannot explain why the level of a spouse's economic resources deters a traditional wife from pursuing personal gain from job advancement but does not deter a traditional husband in a comparable situation. England (1989; England and Kilbourne 1990) suggests that earnings from employment may be a resource that provides wives with *potential* power that they choose not to exercise. Her argument is based on an assumption that women are more altruistic than men and are therefore less inclined to bargain to the margins of their market power. However, introducing an assumption of gender differences in altruism to salvage a social exchange explanation of the tied-stayer phenomenon among husbands and wives with traditional gender-role beliefs fails to account for the behavior of husbands and wives who do not share those beliefs.

The neoclassical and social exchange explanations of gender differences in the tied-stayer phenomenon suffer from the same limitation: they fail to acknowledge that gender-role ideology mediates husbands' and wives' interactions over their work and family roles. Gender-role ideology is a form of "culture-mediated power" that shapes relationships between social groups (Lamont and Wuthnow 1990). The "male provider role" is a symbolic construct that provides one readily available model from which couples can define what Hood (1983, p. 7) calls "the mutually recognized right or authority to exercise power in a given area." That particular model has been the dominant one in the United States for much of the 20th century, but it is not the only model. One alternative that has become increasingly viable as employment opportunities for married women have expanded is the "coprovider role." Whether husbands and wives subscribe to either of these models or some other alternative[13] is determined by a variety of factors, including social background, early socialization, and economic constraints (Thornton, Alwin, and Camburn 1983).

Our findings support the notion that the gender-role ideology a wife and husband subscribe to defines the range of role options they consider and the resources and contingencies that are (and are not) taken into account when deciding role responsibilities. For example, among couples subscribing to the male-provider-role model, disruption of the wife's employment is typically not a relevant parameter in decisions about whether the husband will take advantage of a job opportunity at a new location. Conversely, unless the economic viability of the family is threatened,

[13] See Atkinson and Boles (1984) and Scanzoni et al (1989) for elaboration of other alternatives to the male provider and coprovider models.

relocation for the wife's job advancement is not among the choices considered when couples with traditional gender-role beliefs negotiate role behaviors and responsibilities.

In short, gender-role ideology defines the limits of what is subject to negotiation with respect to work and family roles. Thus, traditional wives' subordination of their job interests to those of their husbands is not the outcome of an explicit, zero-sum bargaining process in which traditional husbands have the resources to make their private interests prevail. Instead, a traditional husband's power is indirect and culturally mediated (Lamont and Wuthnow 1990) to the extent that his role as provider is taken for granted and mutually recognized as legitimate by both spouses.

Empirically demonstrating the effect of indirect or culturally mediated power can be difficult, since it may not be played out explicitly in micro-level negotiations between husbands and wives. Indeed, husbands and wives often agree on the fairness of their decisions about work and family roles regardless of the relative resources each brings to the decision-making process (Godwin and Scanzoni 1989*b;* Sexton and Perlman 1989). Our approach has been to show the consequences of culturally mediated power through asymmetries in the determinants of the tied-stayer phenomenon among those subscribing to different cultural notions of husbands' and wives' work roles. We have detected strong asymmetry consistent with differences in the content of respondents' gender-role beliefs. Among those with traditional beliefs, the wife's potential loss in the job market as a tied mover is not a consideration when the husband is faced with a job offer elsewhere, but it is for those with nontraditional beliefs.

Gender-Role Ideology, Family Migration, and Social Change

Recent research on the causes and consequences of changing gender-role beliefs allows us to speculate on the dynamics of change regarding family migration decisions. For example, we know that cohort succession is an important source of attitude change and that female labor-force participation is both a cause and consequence of progressive gender-role beliefs (Mason, Czajka, and Arber 1976; Thornton et al. 1983; Thornton 1989). According to our measure, in 1977, 25% of the men and 43% of the women in dual-earner couples held nontraditional gender-role beliefs. No doubt, the percentage subscribing to such beliefs has increased since then, and over the same period female labor-force participation has continued to rise, while the gender gap in wages has closed modestly.

Together, these trends suggest that the tied-stayer phenomenon affects men and women more equally today than it did in 1977. A higher proportion of men are likely to be in relationships in which the spouse's employ-

ment is an important consideration when deciding whether to move for job advancement. Moreover, as wives' job- and firm-specific investments have increased, it has probably become more difficult for couples to maintain that a move for the husband's job advancement is in the economic interest of the entire family. As a result, it is probably increasingly the case that geographic locations that are optimal in terms of economic gain for the family as a whole are suboptimal from the perspective of either spouse individually.

In short, compromise over geographic relocation for job advancement is probably a more salient issue among dual-earner couples today than it was in the 1970s. It may also be that the range of culturally acceptable accommodations is changing as well. For example, delayed marriage, separate residences, and long commutes may increasingly become part of the constellation of ideologically sanctioned gender-role arrangements of the 1990s.

The above speculations are based on the premise that males with non-traditional gender-role beliefs who are evaluating job opportunities in a new location are in fact sensitive to disruption of their spouses' employment. While our results for the 1970s showed this to be the case, they also showed that nontraditional males were not nearly as sensitive to their spouses' job circumstances as were nontraditional females. Thus, to the extent that husbands continue to devalue wives' roles as coproviders, women in dual-earner couples will remain disproportionately represented among tied stayers.

Directions for Future Research

We are confident that our findings regarding gender-role beliefs and the asymmetries in the tied-stayer phenomenon are robust and replicable. However, our research was based on secondary analysis and is limited by the information available in an existing data set not explicitly designed to study family migration decisions. A more definitive study would: (1) interview both the husband and wife in each household to obtain complete information on each spouse's work and family investments and potential losses from relocation, (2) pose hypothetical options that explicitly specify economic gains from a move, and (3) follow actual job changes and geographic relocations with longitudinal data.

More important, our secondary analysis did not allow us to study the process by which husbands and wives negotiate their role responsibilities and the way in which gender-role ideology shapes that process. To complement the survey methodology relied on here, couples could be asked to discuss how they would respond to a job opportunity for one spouse in

a new location. The analyst could study the resources that are considered relevant to each spouse, which costs and benefits are weighed and which ignored and the range of alternative outcomes that the couple considers, as well as the couple's report of which outcome they would be likely to choose. A variation on this methodology that has been used successfully by others (e.g., Godwin and Scanzoni 1989a, 1989b) would be to have couples recollect and report on actual decision-making episodes regarding geographic location and job advancement. A research agenda along these lines is likely to provide important insights into the economic, structural, and cultural factors that shape how individuals in families reconcile their personal interests in job advancement with others' interests and with overall family well-being.

APPENDIX

TABLE A1

DESCRIPTIONS OF INDEPENDENT VARIABLES

Variable	Metric	Description
Job- and firm-specific investments:		
Job security	1–4	Average of two items. (1) "The job security is good" and (2) "How likely is it that during the next couple of years you will lose your present job and have to look for a job with another employer?"
Stake in job	1–4	"I have too much stake in my job to change jobs now."
Skill utilization	1–4	"My job lets me use my skills and abilities."
Specific training	Years	Specific vocational preparation (SVP): from the *Dictionary of Occupational Titles* (4th ed.), SVP is "the amount of time required to learn the techniques, acquire the information, and develop the facility needed for average performance in a specific job-worker situation" (Roos and Price 1981, p. 9). Each respondent is assigned the mean SVP score for his or her three-digit census occupation code, as computed by Roos and Price (1981).
Firm size	Employees	Number of employees at establishment where respondent works.
Self-employed	Binary	Primary employment from own business.
Intrinsic rewards	1–4	Average of three items (1) "What I do at work is more important to me than the money I earn"; (2) "The work I do is meaningful to me"; and (3) "The work is interesting."
Earnings	Dollars	Total compensation from primary job, before taxes and deductions. Includes overtime pay, bonuses, commissions, and the like.

1264

General skills:

Part-time work ...	Binary	1 if respondent works 30 hours or less per week.
Part-year work	Binary	1 if respondent works 44 or fewer weeks.
Work discontinuity	Proportion	Proportion of years not worked since age 16.
Gender composition of occupation	Proportion	Proportion female in respondent's three-digit occupation, based on April 1971 CPS
Laborer...	Binary	Respondent's three-digit census occupation code is in the major census group of "laborers, except farm."
Education....	Years	Years, assigned to survey category midpoints.
Age..	Years	Age at time of survey

Spouse's job investments:

Spouse's earnings	Dollars	Spouse's earnings, before taxes and other deductions.
Spouse's specific training	Years	Mean SVP score of spouse's three-digit occupation (see "Specific training," above).

Family investments

Any children in household....	Binary	Whether at least one child lives in household.
Responsibility for a child	Binary	"If someone has to be home with your child(ren) or do something for (him/her/them) when you are both supposed to be working, which of you is more likely to stay home?" (1 if respondent; 0 if spouse, "it depends," spouse not in labor force, or childless.)
Time married	Years	Years married to current spouse.

Gender-role ideology:

Nontraditional beliefs....	Binary	1 if respondent (1) "disagrees" or "strongly disagrees" with the statement, "It is much better for everyone involved if the man earns the money and the woman takes care of the home and children"; *and* (2) "agrees" or "strongly agrees" with the statement, "A mother who works outside the home can have just as good a relationshp with her children as a woman who does not work."

REFERENCES

Atkinson, M. P., and J. Boles. 1984. "WASP (Wives as Senior Partners)." *Journal of Marriage and the Family* 46:861–70.
Becker, Gary S 1974 "A Theory of Marriage " Pp 299–344 in *The Economics of the Family: Marriage, Children, and Human Capital*, edited by T. W. Schultz. Chicago: University of Chicago Press.
Bernard, Jessie. 1981. "The Good Provider Role: Its Rise and Fall." *American Psychologist* 36:1–12.
Blood, Robert O., Jr., and Donald M. Wolf. 1960. *Husbands and Wives*. Glencoe, Ill.. Free Press.
Blumstein, Paul, and Pepper Schwartz. 1983. *American Couples*. New York: Morrow.
Brown, Charles, James Hamilton, and James Medoff. 1990. *Employers Large and Small* Cambridge, Mass.: Harvard University Press.
Da Vanzo, Julie. 1976. "Why Families Move: A Model of Geographic Mobility of Married Couples." Working paper no. R-1972-DOL. Santa Monica, Calif.: RAND.
Duncan, R. Paul, and Carolyn Cummings Perrucci. 1976. "Dual Occupation Families and Migration." *American Sociological Review* 41:252–61.
Emerson, Robert M. 1976. "Social Exchange Theory." *Annual Review of Sociology* 2:335–62.
England, Paula. 1982. "The Failure of Human Capital Theory to Explain Occupational Sex Segregation." *Journal of Human Resources* 17.358–70.
———. 1984. "Wage Appreciation and Depreciation. A Test of Neoclassical Economic Explanations of Occupational Sex Segregation." *Social Forces* 62:726–49.
———. 1989. "A Feminist Critique of Rational-Choice Theories. Implications for Sociology." *American Sociologist* 20:14–28.
England, Paula, and George Farkas. 1986. *Households, Employment, and Gender: A Social, Economic, and Demographic View*. New York. Aldine.
England, Paula, and Barbara Stanek Kilbourne. 1990. "Markets, Marriages, and Other Mates: The Problem of Power." Pp. 163–88 in *Beyond the Marketplace. Rethinking Economy and Society*, edited by Roger O. Friedland and Alexander F. Robertson New York. Aldine.
Godwin, Deborah D., and John Scanzoni. 1989*a*. "Couple Consensus during Marital Joint Decision-Making: A Context, Process, Outcome Model." *Journal of Marriage and the Family* 51:943–56.
———. 1989*b*. "Couple Decision Making: Commonalities and Differences across Issues and Spouses." *Journal of Family Issues* 10:291–310.
Haas, Linda. 1986. "Wives' Orientation towards Breadwinning: Sweden and the United States." *Journal of Family Issues* 7:358–81.
Hardesty, Constance, and Janet Bokemeier. 1989 "Finding Time and Making Do: Distribution of Household Labor in Nonmetropolitan Marriages." *Journal of Marriage and the Family* 51:253–67.
Hood, Jane C. 1983. *Becoming a Two-Job Family*. New York: Praeger.
Lamont, Michele, and Robert Wuthnow. 1990. "Betwixt and Between. Recent Cultural Sociology in Europe and the United States." Pp. 287–315 in *Frontiers of Social Theory*, edited by George Ritzer. New York: Columbia University Press.
Lichter, Daniel T. 1980. "Household Migration and the Labor Market Position of Married Women." *Social Science Research* 9.83–97.
——— 1983. "Socioeconomic Returns to Migration among Married Women." *Social Forces* 62:487–503.
Long, Larry H. 1974. "Women's Labor Force Participation and the Residential Mobility of Families." *Social Forces* 52:342–48.
Markham, William T., Patrick O. Macken, Charles M. Bonjean, and Judy Corder.

1983. "A Note on Sex, Geographic Mobility, and Career Advancement." *Social Forces* 61:1138–46.

Markham, William T., and Joseph H. Pleck. 1986. "Sex and Willingness to Move for Occupational Advancement. Some National Sample Results." *Sociological Quarterly* 27.121–43.

Mason, Karen O., John L. Czajka, and Sara Arber. 1976 "Change in U.S. Women's Sex-Role Attitudes, 1964–1974." *American Sociological Review* 41:573–96.

Maxwell, Nan L. 1988. "Economic Returns to Migration: Marital Status and Gender Differences." *Social Science Quarterly* 69:109–21

McDonald, G. W. 1980. "Family Power: The Assessment of a Decade of Theory and Research, 1970–1979." *Journal of Marriage and the Family* 42:841–54.

Mincer, Jacob. 1978. "Family Migration Decisions." *Journal of Political Economy* 86:749–75.

Perry-Jenkins, Maureen, and Ann C. Crouter. "Men's Provider-Role Attitudes: Implications for Household Work and Marital Satisfaction." *Journal of Family Issues* 11·136–56.

Polachek, Solomon. 1976. "Occupational Segregation: An Alternative Hypothesis." *Journal of Contemporary Business* 5:1–12.

Quinn, Robert P., and Graham Staines. 1979. *Quality of Employment Survey, 1977: Cross Section.* Ann Arbor, Mich.: Inter-University Consortium for Political and Social Research.

Raush, Harold L. 1977. "Orientations to the Close Relationship." Pp. 163–88 in *Close Relationships: Perspectives on the Meaning of Intimacy,* edited by George Levinger and Harold L. Raush. Amherst: University of Massachusetts Press.

Roos, Patricia A., and June Price. 1981. "Fourth Edition Dictionary of Occupational Titles (DOT) Scores for 1970 Census Categories. Codebook." Washington, D C.: National Academy of Sciences, Committee on Occupational Classification and Analysis (mimeo).

Sandell, Steven H. 1977. "Women and the Economics of Family Migration." *Review of Economics and Statistics* 59.406–14.

Scanzoni, John. 1970. *Opportunity and the Family.* New York· Macmillan.

Scanzoni, John, Karen Polonko, Jay Teachman, and Linda Thompson 1989. *The Sexual Bond: Rethinking Families and Close Relationships.* Newbury Park, Calif.: Sage.

Sexton, Constance C., and Daniel S. Perlman. 1989. "Couples' Career Orientation, Gender Role Orientation, and Perceived Equity as Determinants of Marital Power." *Journal of Marriage and the Family* 51:933–41.

Spitze, Glenna. 1984. "The Effects of Family Migration on Wives' Employment: How Long Does It Last?" *Social Science Quarterly* 65:21–36.

Szinovacz, Maximiliane E. 1987 "Power." Pp. 651–94 in *Handbook of Marriage and the Family,* edited by Marvin B. Sussman and Suzanne K. Steinmetz. New York: Plenum.

Thompson, Linda, and Alexis J. Walker. 1989. "Gender in Families: Women and Men in Marriage, Work, and Parenthood." *Journal of Marriage and the Family* 51:845–71.

Thornton, Arland. 1989. "Changing Attitudes and Family Issues in the United States." *Journal of Marriage and the Family* 51:873–93.

Thornton, Arland, Duane F. Alwin, and Donald Camburn. 1983. "Causes and Consequences of Sex-Role Attitudes and Attitude Change." *American Sociological Review* 48·211–27.

Winship, Christopher, and Robert D Mare. 1984. "Regression Models with Ordinal Variables." *American Sociological Review* 49:512–25.

Young, M., and P. Willmott. 1973. *The Symmetrical Family.* New York: Routledge & Kegan Paul.

The Culture of Production: Aesthetic Choices and Constraints in Culinary Work[1]

Gary Alan Fine
University of Georgia

The creation of objects of "aesthetic value" is not merely a topic of philosophical speculation, but is a distinctly sociological activity. Each occupation maintains a sense of superior production (an "occupational aesthetic") that is not reducible to organizational demands. This perspective extends the production of culture approach that sees art as being like all work, suggesting, in contrast, that all work is like art. An aesthetic component to work is reflected in the desire to produce objects (or perform tasks) so as to demonstrate the competence of the worker, as exemplified in a case study of work in four restaurant kitchens. The production of quality is not unbounded, as client demands, organizational efficiency, and the organization's resource base have effects. The centrality of an aesthetic orientation depends upon the market niche of one's organization, career stage in the occupation, and the nature of the work task.

De gustibus non disputandum. [Latin proverb]

How is "good" work possible, given demands for autonomy and organizational constraints on that autonomy? Unfortunately sociologists of work have been little concerned with how work gets done, as that doing relates to questions of style and form: the aesthetics of work. We have lost sight of the conditions that produce "quality," while emphasizing the technical, functional, and goal-directed doings of workers and how workers attempt to undercut authority in the workplace. This choice means that we often examine work worlds from the outside, little realiz-

ing that what is useful to the consumer may (or may not) be elegant to the worker. We do not examine ways in which organizations facilitate and restrain occupational aesthetics. This lack of theoretical interest in form and content may be excused in studies of occupations we label "industrial" or "professional," but it is curious that this deemphasis on the sensory components of work occurs in studies of occupations that involve aesthetic production.

Extending a "production of culture" approach (Peterson 1979; Becker 1974; Hirsch 1972), which analyzes cultural production by the same tools as industrial work, I argue that (1) issues of quality are central to production and that process involves "aesthetic choices," (2) aesthetic choices are a form of organizational decisions, are capable of being negotiated, and are not fully reducible to organization demands, (3) organizational features encourage, channel, and limit explicitly aesthetic choices, and (4) organizations can define their own aesthetics, given their placement within a market niche and clients' definitions. There has been a tendency in sociology of culture (see Wolff 1983) to downplay aesthetic choices, effects, and constraints. My goal is to demonstrate how options and constraints produce the expressive form of work products: what we might term the *culture of production*. I hope to demonstrate how organizational, market, and client constraints affect the qualities of work products.

In speaking of the expressive side of production, I select the slippery term "aesthetics" to refer to the sensory component of production.[2] Why aesthetics? This concept is the broadest of a cluster of terms that involve the sensory qualities of experience and objects: beauty, creativity, elegance, goodness, and the like. For purposes of this analysis, an aesthetic object (or act) is defined as an object (or act) that is intended to produce a sensory response in an audience (e.g., Shepard 1987; Wolff 1983). No special brief other than its utility and general reasonableness exists for this definition. It captures the cognitive (satisfaction) and affective (sensory) components of aesthetic judgments, and also includes the intentional quality of human action. Aesthetics reminds us that these choices are distinct from purely instrumental and efficient choices: workers care

---

[2] The study of aesthetics has been filled with conflicting assumptions and opinions. Philosophers rarely choose to examine situations in which aesthetic decisions are made in the messy reality of everyday life and suggest that aesthetic judgments transcend the production of an aesthetic object and its socially situated character (e g., Diffey 1984; Hincks 1984) These explanations, focusing on qualities of mind (Aldrich 1966; Stolnitz 1960) or the qualities of an object (Beardsley 1958) that produce the recognition that one has had an aesthetic experience (Wolff 1983; Shepard 1987), downplay the sociological interest in the interactional, relational, or institutional features of aesthetic evaluation (see Dickie 1974; Danto 1981).

about "style," and not only about technical quality. Although form and function are typically intertwined, aesthetics refers specifically to the production of form, not only to function. Attempts to produce "good work" often involve an intimate linkage between form and function, and functionally perfect objects may be seen as having perfect form. Judgments of quality adhere to both form and function, although the focus here is on the former. In cooking, and other work arenas, the sensory characteristics of objects (and services) have a special standing in appreciation both among workers and publics.[3]

Sociologists recognize that the practical creation of industrial objects is a fundamentally social enterprise, constructed through interaction and organizational constraint.[4] Yet, the feeling for form or creative impulse, as well as its limitations, needs to be emphasized in theorizing on the structure of work and occupations. Not doing so gives a distorted picture of the workplace, making it alternatively seem too instrumental (denying a sense of identity and craft to workers) or too filled with conflict (emphasizing how workers are separated from their work and their supervisors). Work matters to workers, and workers have craft standards by which they judge work products and performance that transcend the narrow goals of producing things with efficiency and to bureaucratic specification. This connection between the worker and the work is central to the occupational identity of workers. Craft is a part of all work life.

I examine a single occupation, professional cooking, hoping to demonstrate four things. First, cooking, like all occupations, involves an aesthetic concern, which takes its form in decisions about the sensory components of food. Second, the practical doing of cooking is an everyday accomplishment and must be negotiated in practice by workers. Third, culinary production is channeled by social and economic constraints and by occupational segmentation. Finally, this argument is generalized to

---

[3] In this article I bracket the origin of aesthetic choices, wishing to see how such choices are constrained and utilized. My concern is not to trace the dynamics by which particular judgments come to be seen as aesthetic (see Fine 1989), but only those choices that have been accepted by a group of workers Nor am I concerned with the qualities of the object involved. Griswold (1986) argues that the aesthetic involves both elegance (simplicity) and beauty (amplitude) to produce a response. While I use Griswold's distinction to focus on the characteristics of objects, my definition emphasizes the relationship between actors and objects.

[4] Sociologists of aesthetics interested in comparative research must confront two basic presuppositions: (1) that all occupations have aesthetic components, that is, that sensory issues are a part of all work, and (2) that occupations vary on the self-consciousness and centrality of these aesthetic issues to the work. Because this research is grounded on a single occupational case study, I can do no more than suggest the plausibility of these claims.

other occupations, suggesting an integration of the sociologies of work and culture.

THE WORLD OF RESTAURANTS

Whatever cooks may wish to think of their own work, restaurant managers often refer to this economic segment as "the hospitality industry." This phrase reminds us that restaurants are industrial organizations operated for profit by capitalists. Although food must look, smell, taste, and feel good to maintain an audience, this is not sufficient. Food must be priced to be profitable and must be produced consistently and efficiently. Among the techniques used by restaurant managers to achieve profit are paying low wages, hiring few employees, and procuring inexpensive raw materials and equipment. Food services are caught between the demands of aesthetic creation and the viselike grip of free-market capitalism.

Cooks suffer the strains of a set of conflicting ideologies that push them to be artists, professionals, businessmen, and manual laborers (Fine 1982). Because professional cooking is situated amid demands for aesthetic choices, consistency, efficiency, autonomy, and highly skilled technical work (Hall 1975, pp. 188–200), it provides a challenging site from which to examine the development, conflicts, and negotiations of sensory judgments at work. Whyte (1948; Gross 1958) notes that restaurants are both production and service units, providing the cook with two separate "authorities"—managers and customers—adding further strains. Whereas factory workers, beauticians, and sculptors do not have the same balance of concerns, some of the same dilemmas are also present in these work worlds—all occupations combine expressive and instrumental demands, personal freedom, and organizational control in varying degrees.

I conducted participant observation in four restaurants in the Twin Cities metropolitan area (St. Paul/Minneapolis and their surrounding suburbs), spending a month observing and taking notes in the kitchen of each restaurant during all periods in which the restaurant was open.[5] In each restaurant I interviewed all full-time cooks—a total of 30 interviews. Interviews lasted approximately 90 minutes, with some over three hours long. Field notes and interviews from this material are identified throughout the text.

The four restaurants provide a reasonable range of professional cooking environments in the Twin Cities. These four restaurants are not a representative sample of all restaurants but represent the upper portion

[5] This constituted approximately 75–100 hours per restaurant.

of Minnesota restaurants in status; they are not "family," "fast food," or "ethnic" restaurants:

1. La Pomme de Terre is an haute cuisine French restaurant, by all accounts one of the best and most innovative restaurants in the upper Midwest.
2. The Owl's Nest is a continental style restaurant, best known for the quality of its fresh fish. Its primary clientele is businessmen, and the restaurant is a multiyear Holiday Award winner.
3. Stan's Steakhouse is a popular neighborhood restaurant located in a middle-class area, not known for the quality of its restaurants It has received metropolitan awards for the quality of its beef.
4. The Blakemore Hotel is part of a chain of hotels that is not esteemed for the quality of its cuisine. The hotel is modern, catering especially to business travelers. The hotel has a banquet service and operates a coffeeshop and hotel dining room.

The restaurants vary widely in the number of customers served—from 500 on a busy weekend evening at Stan's to about 75 on the same evening at La Pomme de Terre—but each hires from five to 10 cooks of whom usually three or four are at work at any one time.

Although there is not space for a full ethnography of restaurant life (see Fine [1987]), the large majority of cooks (80%) observed were male; most were in their twenties. While the background and training of the cooks varied considerably—individually and by restaurant—each of the head chefs was trained at a local technical-vocational institute, and many of their assistants were similarly trained (for details of this training see Fine [1985]). This was especially true at La Pomme de Terre and the Owl's Nest, whereas at Stan's Steakhouse many cooks were promoted from dishwashers and most did not see cooking as a long-term occupation.

DOING AESTHETICS

All work is socially situated and constrained environmentally and organizationally. No matter how idealistic the goals of the worker, ultimately these goals are embedded in the negotiated compromises of work. Howard Becker, discussing art as work, claims that aesthetics is ultimately activity rather than a doctrine (Becker 1982, p. 131)—it is an everyday accomplishment. Theory only flickers around the edges of the consciousness of workers. It follows from this that most workers are not explicit about (or even conscious of) their aesthetic decisions.[6] They desire to

[6] Although all cooks have aesthetic orientations, most do not fit in that category of workers that we describe as artists. For an occupation to be an art world, it requires (1) that a group of persons be working toward a common end. The group should be

produce objects or services that are pleasing sensually, but typically the basis on which they realize this is vague. For example, a hotel cook told me: "When I make my soup . . . I try to make it look as nice as possible, and to taste. I feel I take a lot of pride in it. When other people make soup it doesn't always look like mine" (field notes, Blakemore Hotel). This worker has a generalized sense of "niceness" that includes looks and taste, but analysis does not transcend this partially inarticulate sentiment (Fine 1987).

The content of this sensibility varies by cook and restaurant and is further complicated by the realization that cooking involves situated choices. Still, all cooks hope to present what they consider appealing dishes of which they are proud—food that will appeal to their customers' senses, not merely food that will satiate them or make them healthy (the functional characteristics of food). This culinary evaluation involves numerous senses. The head chef at La Pomme de Terre responded when I asked what he liked best about cooking:

> Making something that I think is just the greatest. I did a bouillabaisse . . . and I thought it was just the greatest. . . . It had a lot of seafood in it, a lot of shellfish, shrimp, lobster, mussels, clams, and about six other seafood items in it, and the sauce was a somewhat thin, primarily lobster-based sauce, lots of butter, and very, very rich, and the thing that was best about it was everything was made to where, typically if you have bouillabaisse, you have to hold onto something with the tongs and dig meat out of the shell and stuff like that, but I prepared it so that everything was done for you. . . . It was not only tasty and unusually fantastic as far as flavor, smell, and sight; it was easy to eat. (Interview, La Pomme de Terre)

The range of senses is implicated in this cook's sense of his culinary triumph.[7] Lest one believe that this sensory concern applies only to those finer restaurants (where some might claim the cooks really are artists), it applies to the steakhouse as well. The chef at the steakhouse responded

---

aware of each other and should have social and professional contacts, (2) that the group have an artistic theory to guide them and to demonstrate their shared commitment. A theory of art is essential (see Danto 1964, p 581); and (3) that there be a set of recognized institutional gatekeepers and gatekeeping organizations that choose candidates for ascension into the canons of art—what Dickie (1974) refers to as the "institution of art." Cooks in the Twin Cities lack tight networks, a widely held aesthetic theory, and acceptance by artistic gatekeepers. While one might discover a culinary art world in small sectors of the hospitality industry in New York, New Orleans, or San Francisco, throughout most of the rest of the country, cooks just cook.

[7] "Occupational triumphs" consist of occasions in which workers feel that they have operated to the limits of their jobs—they are "pushing the envelope." Working within the rules, they have transcended them, demonstrating in their own minds at least that they are not mere workers, but true artists, true professionals, or the like. They have produced not just an object but a memory that they can narrate to convince others of their virtues, even given constraints and normal operating procedures

to my question about what a piece of baked salmon should be: "It should be just very lightly, you should see a tinge of brown on the outside, but it shouldn't be overcooked. It should be just done. Nice and moist" (interview, Stan's). Again, a range of sensory modalities affects the evaluation of food, even where one might assume that such interest is limited.

Evaluation need not only involve the production of sensory appealing products, but may also adhere in the sense of doing—an experience that we might liken to that of "flow" (Csikszentmihalyi 1975). Here the doing is the end. Some cooks speak of themselves in terms of their actions (Clark 1975, p. 33), making cooking into a performance art:

> It's very much like an actor preparing to go on stage and go into work and start in a quiet pace and figure what you're going to be doing; you get your equipment ready, sharpen knives, cut meats, trim your fish, and make your vegetables and make your sauces and get everything set up and it gets a little bit hotter; people start talking more and the waiters start coming in, and this is going on over here, and by the time everything starts coming together, it's like you're ready to go onstage. It's there. . . . Once the curtain goes up, everyone knows exactly what they're supposed to do. (Interview, La Pomme de Terre)

For some, the criteria for quality labor are primarily in the product (the sight, feel, taste, or smell), for others they are in the performance, but for each, the work has a style, a sense of form, an aesthetic.

Ideally this evaluation should be grounded within the occupation—although products are typically also judged by clients and on occasion performance is as well (as in the proliferating demonstration kitchens). The evaluation of production is not only a function of demands of customers and managers; cooks see themselves as having independent standards of judgment. Certainly these independent standards cannot radically vary from the demands of their customers, even for elite chefs (Kimball 1985, p. 18), and there are critical situations in which clients' demands take precedence, but cooks have their own judgments that are not reducible to organizational requirements. Management and customers do demand aesthetic production, and, so are in sympathy with the goals of the cooks, but the constraints that they demand and their standards of aesthetics may limit what cooks are able to produce. All parties want good work, but the meanings and the external considerations differ.

The salience of evaluations by cooks is evident when workers are creating "unique" items. This follows from the observation that the more special the product and the less routine the task, the less an organization can rely on formal rules, and the greater the autonomy that must be given to workers (e.g., Woodward 1965; Faulkner 1971; Coser, Kadushin, and Powell 1982). Individualized production technologies lead to choices, but can also simultaneously (as I shall describe later) lead to a

recognition of the lack of autonomy from constraints. When cooks can create without pressure, they do, and are proud of the results. For example, one cook, preparing a wedding dinner, carved a pair of birds of paradise from apples and sent them to the bride and groom as *his* gift for their marriage (personal communication, Robert Pankin, 1987). Likewise, after making a chocolate cake, the pastry chef at La Pomme de Terre added four raspberries and drizzled chocolate sauce over them, commenting, "I'll put some fruit on here so it looks a little more abstract" (field notes, La Pomme de Terre). Her touch was not a result of management policy (although she was expected to make "beautiful" desserts); rather, the standards and techniques she used developed out of her sense of what it meant to be a competent pastry chef.

Although cooks have some measure of control over the sensory characteristics of the food they prepare, the doing of this aesthetic work is an everyday achievement; it is not merely grounded in theoretical choices. The production of "high quality" items, as defined by cooks, depends on a balance of culinary ideals (e.g., using natural ingredients) and production constraints. The ends direct production choices, as in two separate discussions of the color of a sauce:

> The head chef at the Owl's Nest pours a considerable amount of Gravy Bouquet in his Brown Sauce to make it "richer." He then adds white pepper and stirs the sauce. He tells me that: "Black pepper shows up and looks like mouse turds. Little black specks. So I use white pepper." White pepper is also added to the restaurant's mashed potatoes (Field notes, Owl's Nest)

> The head day cook is preparing cheese sauce, using powdered cheese. He adds a capful of orange food color to the pot, saying that this makes the sauce look more like cheese, and, if you were actually to add cheese, "it gets too sandy." (Field notes, Owl's Nest)

These cooks are making decisions in practice. They believe, certainly correctly, that the visual appeal of the food, the first thing that both cooks and customers notice, affects the way the dish tastes—sensory realms are interconnected (e.g., Moir 1936; Pangborn 1960).[8]

Cooks can be admiring or critical toward what they prepare, based on their evaluation of the outcomes, both instrumentally (success in sales and customer appreciation) and in terms of their occupational standards. This evaluation implies a realm of objects that are considered lacking in

---

[8] For a more extensive analysis of aesthetic ideologies in food preparation see Fine (1985). In practice, cooks negotiate which sensory realm is most significant for particular dishes, but the visual appeal of a dish is typically given greatest weight. This may be because the visual realm is the first encountered or because it is in this sensory domain that schoolchildren are more extensively trained—art classes typically lack stoves or perfume atomizers.

these components that other objects have. No occupational world can long survive if participants judge everything equal to everything else.

For collective judgment, differentiation in the evaluation of produced objects is essential.[9] In cooking this judgment may involve any of the relevant senses. For example, one cook criticized a bunch of grapes as having "bad lines." An outsider might be confused how grapes can have bad lines, until it is learned that the ideal of a bunch of grapes is a pyramid and that other bunches meet this criterion better. Crepes can be described as "lopsided," implying agreement that crepes should be circular. A more detailed example is the condemnation of a particular dish that "doesn't work":

> Howie and Tim taste the beet fettucini that they had planned to serve with a tomato sauce—an orange-red sauce on top of a crimson pasta. Tim says to Howie: "There's something that didn't work. It looks like puke." Howie adds: "It tastes like Chef Boyardee. It tastes like Spaghetti O's. It tastes like snot rag." They decide not to add the sauce. (Field notes, La Pomme de Terre)

This judgment is predicated on their view of what constitutes proper food presentation—which colors go together and what the taste and texture of a properly made sauce should be. Such standards, while based within the occupation, must be echoed by at least some customers. Although the judgments of cooks are never far from their sense of the customers in their market niche, when being creative they use themselves as guides:

> You have the idea in your mind of how something should come out and you have to use your hands and eyes and taste and nose. You have to make it come out the way . . . you want it. (Interview, Stan's)

> The thing is to just have the guts to go in and do it. Just try it. Not worry about is this thing going to work or not. . . . It's color, flavor, texture, smell. It's all those things put together and somehow I have a sense of organizing these things and putting them together. (Interview, La Pomme de Terre)

Cooks do not discuss these judgments in terms of their customers, but in terms of what they believe works, even if they lack a formal theory of what they are doing (Sclafani 1979). There is a set of aesthetic conventions that are based on occupational standards (Becker 1982), separable from organizational demands, but which must be fitted into the constraints imposed (or believed to be imposed) by external sources and by

[9] My argument is that Kant's idea of free judgments of taste is unlikely to be made in most practical aesthetic worlds; rather, aesthetic judgments have a relational character. We judge things in relationship or in comparison with other objects. At some level we are deciding, not whether something is good, but whether it is good of its kind (Kant 1952; Shepard 1987)

the structure of the occupation itself. Occupations struggle to gain control over criteria for judgment from regulators, employers, and clients. Although the recognition of this struggle has been a staple of the analysis of "professions" and other occupations, it applies equally to the control of the aesthetic choices in work.

CONSTRAINTS AND NEGOTIATIONS

Given claims of independence within an occupation, on the one hand, and structural limits, on the other, how do workers produce objects that they consider satisfying and of high quality? What are the dimensions that channel how workers do good work? In order to examine this question, I describe three forces external to occupational autonomy that constrain production choices and show how workers cope with these constraints. In cooking, as elsewhere, organizational constraints not only determine the products but ultimately shape the values of workers. On some occasions, cooks chafe under the restrictions of the workplace, but often these restrictions are taken for granted and treated as merely a reality of the occupation.

Cooking, like all occupations, as Anselm Strauss (1978) emphasizes, is grounded in negotiations and compromises. Cooks strive to control the means and circumstances of production, both to make their own day passably pleasant and to permit them to be satisfied with what they produce.[10] The proximal source of constraints is a restaurant management that depends on the loyalty of its customers, and this pressure is filtered through the head chef who is given an annual or monthly budget with which to work. The irony is that for the same reason management also supports and encourages aesthetic presentations, as long as this good work remains profitable. To satisfy management the chef must manipulate the staff to make a profit and to produce good food. At three of the restaurants studied, the chef received a bonus if he operated within the budget. This control is furthered through the internalized acceptance of these economic and temporal constraints by most cooks.

The ultimate dilemma for cooks is the recognition that often they must serve "bad food"—food that they believe is not up to their own standards of quality, but they have no choice.[11] It is difficult to propose rules for when "poorly prepared" food will be recooked—the etcetera rule,

---

[10] Some workers at the hotel kitchen would come in an hour early (without overtime) in order to set themselves up, feeling that the volunteered time would be worthwhile in improving the quality of their production and permitting them a less hassled day.

[11] Comparative data indicate that this is not unique to this scene, as Walker and Guest (1952, p. 60) describe similar attitudes of autoworkers.

which suggests that no complete set of rules can be formulated, is too prominent (Garfinkel 1967)—but the cost of the food, the time for cooking, the pressure in the kitchen, the status of the customer, the conscientiousness and mood of the cooks, what is wrong (and if it can be partially corrected without recooking), and the status of the restaurant all affect the decision. These decisions can be negotiated among the kitchen staff and with management on the spot, but all cooks must recognize that they must serve food that they know is not up to their standards. Cooks shrug when they send substandard food to unknowing customers and respond sarcastically when, at times, servers announce that they were complimented on these dishes.

One cook described her frustration with a rack of lamb: "I've racked some lamb . . . that was just an abortion. It was just awful; I rolled the pastry too thin and the lamb was overcooked and . . . it came out looking not like it was supposed to. That makes me feel bad, even though that's fine and you have to use it. You can't throw it away, but I feel really bad" (interview, La Pomme de Terre).

Cooks are dismayed when serving food of poor quality, and like so many workers, they deny that they really care by turning the offensive food into a sick joke, engraving role distance in their performances:

> The watercress sauce, created for the salmon appetizer, has separated. Tim (the Head Chef) says sarcastically: "Oh, well, they all look like shit. We don't have to worry." Gerry, his co-worker, jokes: "The room's dark." (Field notes, La Pomme de Terre)

Such joking is legitimate in that cooks have other occupational rhetorics than that of artist to rely upon; for that moment they can constitute themselves as manual laborers, as alienated as any. In occupations, such as cooking, that can draw upon several occupational rhetorics,[12] workers can strategically employ these to preserve their self-integrity. They project themselves into the food that they produce, seeing inner qualities in the outcome. When the food does not meet their standards, they must use techniques for backing away from the equation of self and product. The strategic use of rhetoric is one way of coping with the personal tensions of presentation of self. Switching the available metaphors of their work can serve important ends in preserving role distance and in indicating their control (Fine 1982).

Having demonstrated that cooks are limited in their ability to produce

---

[12] I have spoken of cooks as drawing from the rhetorics of business, art, manual labor, and professionalism to define their work and protect their selves (Fine 1982). Other rhetorics such as craft or sales might affect this occupation on certain occasions.

dishes they consider of high quality, I turn to three forces that prevent their achieving their occupational ideals: customer taste (client demands), time (organizational efficiency), and the economics of the restaurant industry (the resource base of the occupation). These three factors cause cooks to compromise their own taste. Through the constraints of production the production of culture model fits into an analysis of aesthetic choices. To be considered problematic, production depends on a recognition of aesthetic options for constraints.

Client Demands

The restaurant cook prepares food for an audience that does not belong to his or her occupation—an audience that may not have the same standards or even be aware of the existence of standards. Yet, both cooks and customers agree that restaurant food should be aesthetic, whether or not they agree on these expressive dimensions.

Because of the power of the market, autonomy is given up to the expectations of one's audience (Arian 1971). As a result of the loss of autonomy, workers may resent those they work for who do not have their standards of quality and competence—not just bosses, whose sin is cynicism, but also clients, who are seen as culpably ignorant.

Unlike such occupations as beauticians, plastic surgeons, and housepainters in which workers negotiate directly with those who ultimately judge them, cooks must rely on their typification of their audience, given their understanding of the restaurant's market niche.[13] Their evaluation is mediated through managers and servers. Those standing beyond the output boundary are not easily known (see Hirsch 1972; Dimaggio 1977). Dishes are cooked for typifications, not persons; yet, it is persons who have the options to complain. Customers can judge the dish, whereas cooks have difficulty judging the customer.

As a consequence, cooks have developed techniques for dealing with the vagaries of customer taste. At the steakhouse and the continental restaurant it was standard procedure to undercook beef slightly. This allowed for correction if the customer wanted the meat more thoroughly cooked. Steaks can never be cooked less. Still, these cooks became annoyed when customers insisted on having their steaks well done. One

[13] Market niches are in part a function of conscious decisions by managers and chefs to capture audiences. In this they create an establishment that will provide an experience that appeals to a potential pool of clients (e.g., Finkelstein 1989; Shelton 1990). In contrast, niches are occasionally carved by customers who discover establishments; then managers must ensure that they continue to meet the desires of these clients.

Friday night at Stan's, a large number of steaks were sent back, to the cooks' frustration:

> One waitress says to the head chef, referring to the customers: "Are those steaks burnt up enough?" The chef responds: "I hope so. I don't want them." Later another cook comments about the evening: "Bunch of ass-holes out there. They don't know what they want." He means that they don't want what he wishes to serve them. (Field notes, Stan's)

The problem is equally relevant at La Pomme de Terre where the canons of nouvelle cuisine emphasize not overcooking the food and spoiling its "natural" taste. These cooks, too, became annoyed when their "per-fectly" cooked dishes (pink duck breast, translucent fish) are returned for additional work. Not only is the cook's ability questioned by the customer, but cooks believe that by accepting the motto, "The customer is always right," they are prostituting themselves,[14] even though they hope that they may eventually educate their customers (see Becker 1963, pp. 79–100). By pleasing the customer, they deny the validity of their standards. The legitimacy of their aesthetic standards is being invalidated by external demands.

Spices and condiments pose a similar problem. The head chef at the Owl's Nest notes: "You season things, but not completely seasoned. The first thing the customer does is see the salt and sprinkle it on, pepper and et cetera. Takes a bite and puts it down and says this has too much salt on it, and take it back. He was the one who put the salt on it; we didn't. So we underseason things. You have to think for the customer. . . . You have to think of everybody's taste" (interview, Owl's Nest). Even if cooks feel that some foods are unappetizing, they must serve them to customers who enjoy them. Further, even though they personally feel that some foods taste "bad" (e.g., fried liver, spinach), they must learn how to cook them in such a way that the customer who likes them will know that they are cooked correctly, that they represent the best professional practice. They must role-play the standards of their clients. This concern for customer taste (and its limits on cooks) is evident at La Pomme de Terre in the selection of fish specials:

> I ask Tim how they select the two fish specials each night. Tim tells me: "We try to have variety. If we have an unusual one, like with peach, we'll have a conventional one, like the monkfish." (Field notes, La Pomme de Terre)

Customer taste is always taken into account, often explicitly, by cooks.

---

[14] This is a problem that is faced by portrait painters who give up their artistic autonomy to the client. The client feels that he or she has the right to determine his or her personal likeness (see Stewart 1988).

This differentiates them from the higher reaches of the fine arts where, rhetorically at least, obeisance to client demands is considered subversive to an artist's occupational standing.

Organizational Efficiency

Organizations are expected to produce a certain number of products or services in a set time period (Lauer 1981). As a result, temporal demands constrain production decisions in restaurant kitchens (Fine 1990). Customers will wait only so long for any dish to be prepared, and cooks have limited time in which they can prepare for dinner, given the size of the staff, affecting what can be served. These temporal constraints suggest why, discomfiting as it may be, when food falls on a dirty counter or floor after being cooked, cooks will wipe or rinse it, and then serve it, with the customer none the wiser. The illusion of quality demands hidden affronts. Since cooking is a backstage occupation, innumerable depredations to the foodstuffs are possible (e.g., Orwell 1933, pp. 80–81). A steak that takes 30 minutes to cook must be served because of customers' temporal expectations; customers would never wait for a "second try." Likewise, if a fillet of fish breaks while being removed from the pan to the plate, the cook will rearrange it as nicely as possible, but still serve it. The production features of the kitchen and, ironically, the demands of the client, permit no alternative.

Time also affects specific tasks in the kitchen, which, although they would make the food more appealing, cannot be tried because of time constraints. One cook explained that he wishes to do a "French cut" on a rack of lamb, but adds "I'd never have the time to do it" (interview, Blakemore Hotel). Likewise, cooks do not have the time to improve poor quality produce:

> Martha (the day cook) says to Doug (the head chef): "The radishes are bad, but I don't have time to clean them up. . . . These look awful." They are dirty, discolored, and misshapen. Doug sorts through them, and throws out a few of the worst ones, and they serve the others. (Field notes, Stan's)

The problem of timing is particularly acute at Stan's Steakhouse, which, of the four, serves the largest number of customers. Often plates are not wiped off if sauce spills. As one cook joked on a busy evening: "I'm going for numbers, not for quality." Although this is not entirely true, it is truer than it might be under ideal circumstances. *Quality* production is a luxury; production is a necessity.

Time constraints apply not only to particular dishes, but to the creation of more elaborate food presentations. As one cook remarked: "To be creative you need time. You can't always have a deadline behind you.

Because when you do, you're in a rush. And then when you're in a rush you tend to fail with the creativity. 'I need this by such and such a time,' and then you start getting out the same old thing" (interview, Blakemore Hotel). The head chef at La Pomme de Terre learned the day before that he must prepare a large press party for his employer. The chef confides to me that despite an impressive menu (sole turban, smoked goose breast with port wine and fruit, goose liver mousse, and duck galantine): "It's not going to be as good as I'd like. I only learned about it today. I'd like to make a grandiose first impression. . . . It's a matter of pride. The artist's pride is at stake" (field notes, La Pomme de Terre).

Ideas for a large display with fresh lobsters and a lobster mousse had to be shelved for lack of time. Although the owner felt that the party was a great success, the head chef was disappointed because it did not measure up to the quality of which he felt they were capable. While the organization was technically efficient—it did produce *something*—it was not sufficiently aesthetically productive, given the aesthetic standards of the chef.

Resource Base

The final constraint is the cost of materials. Cooks must remind themselves that ultimately they are part of corporate capitalism—what Blau (1984, p. 10), studying architects, terms "professional practice." Indeed, in few other market segments does a truly free market operate as clearly as in the restaurant industry.

Price and quality combine together to determine restaurant success, as judged by external publics. Restaurants are known directly by clients who learn about them through advertising, experience, word of mouth, the publicity of managers, and institutionalized gatekeepers, such as critics and journalists. On some fundamental level, price and quality conflict, and the manager and head chef must decide to which market niche to appeal, given their perception of the organizational ecology. The head chef at the Owl's Nest recognized these economic trade-offs: "We always have variables. The compromise in your mind is using the best you can use, and still putting it into an affordable level for the average customer" (interview, Owl's Nest). As decisions are locally situated, this trade-off involves specific decisions about particular products, rather than an absolute rule of thumb:

> In theory the head chef of the Owl's Nest believes in using the best that is available. He explains· "The customer may not be able to tell in the finished product. The finished product might taste the same, but it should be made that way." However, when I ask later why he adds cheap Ameri-

can cooking wine to sauces, rather than expensive French wine, he claims·
"People can't tell the difference." (Field notes, Owl's Nest)

Of course, the question is best for what? Imported truffles, beluga caviar, and Chateau Margaux add enormously to the cost, but only slightly to the taste. For this chef the possibility of adding these other expensive ingredients is not even a part of his consideration, until a sociologist brings them up. The economic reality of food preparation affects his aesthetic vision.

According to the staff at La Pomme de Terre, what distinguishes them from elite American restaurants is not the quality of the preparations, but "the touches"—those extra garnishes that restaurants can afford to add if they have a large staff and a loyal clientele. They compare their restaurant to others of which they are aware, and find themselves wanting:

> The owner confides to me that one of the Twin Cities restaurant critics said that La Pomme de Terre was the best restaurant in the Twin Cities, but not as good as Le Perroquet (Chicago) or Lutece (New York). He explains: "I asked him why. He said, 'The touches ' . . . They have more people in the kitchen. The difference is volume They can count on being sold-out every night of the week. We can't." (Field notes, La Pomme de Terre)

Timing, customer taste, and resources merge to prevent this restaurant from reaching its potential, as filtered through the owner's estimation of the Twin Cities restaurant market. A year later, this man opened a restaurant that was more expensive and formal than La Pomme de Terre, and included "the touches." It failed; the market was not there. Cultural products have different price elasticity, even within particular niches. Some food prices are simply considered "obscene." There is an obdurate reality that prevents unconstrained aesthetic activity.

As a consequence, cost must be considered by decision makers. The staff at La Pomme de Terre experimented with different blends of coffee to find a mix that had the richness of expensive coffee with as much inexpensive coffee as possible. Likewise, the pastry chef commented about a raspberry-lemon gateau: "It's called, 'Let's be creative using the leftovers'" (field notes, La Pomme de Terre).

The skill in running a profitable organization is to provide goods or services that clients desire and that appear to be worth more than they cost to provide. Some foods seem expensive, but are not. When the head chef at La Pomme de Terre created Saffron Pasta with Lobster Sauce, he noted that the food cost "is not all that high." Likewise, the head chef at the Blakemore Hotel explains that salami horns filled with cream cheese look elegant, but are inexpensive.

An ability to compromise on quality when one's judgments conflict with the economics of the organization is crucial for advancement. The head chef at La Pomme de Terre had planned to promote his head day cook to sous (assistant) chef, but decided against it:

> Because he's such a renegade. I can't rely on him to do what I want him to do. . . . As an example, last week he's been doing that veal special that he came up with and it's a real beautiful dish. He takes the veal roulade and he puts prosciutto ham and goat's cheese with herbs and folds it over and sautés it, and serves it with tomato sauce. It's a good dish. He had a couple in there that were getting a little bit dark. The veal starts to get a sort of grey when it gets old, but they were fine; they were just starting to turn grey. I looked at them, and I said they're fine . . . and he was putting up a couple of veal specials, and I went in the walk-in and those suckers were sitting there . . . I called him in, and said, "What is that, for your mother or what? Come on and get moving. This is a restaurant " He's got such a paranoid pride over being criticized for something that he just took it upon himself to do it. . . . He doesn't have the concept that we're in business. He just thinks it's one big happy deal. (Interview, La Pomme de Terre)

This cook placed his standards of quality (standards with which in theory his head chef would agree) above the production needs of the organization, and, being unwilling to negotiate, lost his opportunity for promotion.[15] Cooks must keep one eye on the stove and the other on the marketplace, balancing their sensibilities with what the hospitality industry will permit. While chefs and cooks negotiate with each other, and chefs negotiate with managers as to the boundaries of their decision making and their commitment to quantity and quality (e.g., the number of scallops to serve or the time at which food begins to be "off"), an economic imperative channels the ability to produce.

THE SEGMENTATION OF AESTHETIC WORK

Although each occupation reveals concern with the expressive quality of production, comparative analysis would demonstrate that this concern is variable, not absolute; it certainly is expressed in different forms that may be more or less central to the doing of work. I have argued above that some outcomes and performances are seen by workers as having more value than others. Further, a determination of what constitutes quality is not absolute within an occupation or art world. There is no single aesthetic sense or unified set of conventions. Painters do not paint

[15] This is a story for those with a sentimental attachment to a happy ending: within a few years this young man had become head chef at an outstanding, creative restaurant in the Twin Cities. By then he had learned to control his employer's costs.

alike, and they do not believe that they should. Even the task influences one's orientation to work and the role that aesthetic or sensory concerns should have in production. Every occupation is socially segmented (Bucher 1962), and it is the effect of this segmentation that I wish to explore. Cooking is segmented on several dimensions; three of the most prominent are the restaurant's status, the cook's career stage, and the work task—reflecting differences among organizations, actors, and events.

Restaurant Status

Cooks differ in their working environments—the types of restaurants for which they cook. Freeman and Hannan (1983), detailing the importance of market niche in organizational ecology, focused on the restaurant industry. Restaurants are competitive small businesses in a segmented environment. In this free market, product differentiation is crucial. That restaurants and their cooks have different standings and variable amounts of cultural capital is a function of the market niche to which the restaurant aspires and of the "background culture" of the cook (Fine 1979, 1989).[16] When the cook and the restaurant management do not share a cultural orientation, the cook must cook "up" or "down" to the level of the restaurant: the cook's display of his or her cultural capital becomes a form of impression management.

Some restaurant managers expect cooks to have a sharp sense of sensory or aesthetic issues in their cooking—to be aware of the subtle permutations of smell, taste, texture, and looks—and to use this culinary sense with relative autonomy. Cooks at La Pomme de Terre were more overtly concerned with individual choices than were cooks at the other restaurants, and they were given more autonomy in the expectation that they would be creative. These cooks never looked at recipes; they created new dishes or cooked from memory. The employees of the hotel kitchen and the steakhouse were less self-consciously concerned with the aesthetic quality of their dishes, although they made creative decisions and felt pride in the appearance and taste of their food. Time and motivation in these establishments sometimes led to food being served that might not

---

[16] Bourdieu (1984) uses food consumption in France as an indication of the cultural capital of the eater, but it is also true that food production is an indicator of the cultural capital of the cook We are known by what we eat, but we are also known by what we cook. The more sophisticated cooks, better trained, raised in more sophisticated homes, or more impelled by the goals of their restaurant, are more attuned to the dishes that represent haute cuisine and demonstrate the existence of cultural capital.

have been served elsewhere (e.g., onion rings with breading that was falling off)—they did not have time for elegant and creative production.[17]

The self-image and market niche of a restaurant affects how workers view the sensory qualities of their production. Although McDonald's and Lutece have aesthetics associated with the work,[18] the cooks at the latter have more autonomy and their aesthetic decisions are more subtle and consequential. McDonald's has corporate aesthetic standards for the "design" of their food, set by the central office. Worker aesthetics at McDonald's involves problem solving of immediate production needs— following the preset rules with style, care, efficiency, and coping with customer demands.

Career Stage

A concern with aesthetic issues has different salience at different stages of a cook's career. These stages are often correlated with organizational position because many workers move up the restaurant hierarchy as they demonstrate competency. Jobs change as individuals mature within their occupation and achieve higher status. Different values, goals, and opportunities affect how aesthetic preferences will affect actual production decisions.

Entry-level cooks are often required to perform routine manual labor, unlikely to be defined in terms of aesthetic choices. They may be asked to chop onions, peel potatoes, or destring celery. As they progress through their careers, they are given more responsibility, and with this responsibility comes the authority to know (Mukerji 1976)—to prepare and later to create complex dishes. This responsibility emerges when the cook demonstrates talent, competence, and conscientiousness to his or her supervisors. I asked a junior cook at La Pomme de Terre whether she had created any dishes:

> Cook· I haven't been allowed the freedom to. I think I will.
> Author Is there any dish you want to try?
> Cook: Yeah. I did a rainbow trout stuffed with spinach and mushrooms

[17] The comparative analysis of restaurant aesthetics as a function of organizational goals is an important topic, but one that I lack the space to confront. The market niche of the restaurant and the cultural capital of customers, cooks, and managers affect the aesthetic choices and constraints that channel presentations. The elaboration of dishes is much greater in La Pomme de Terre than in Stan's, which relies on simple, minimally transformed preparation of foodstuffs.

[18] One reader of this paper commented that the aesthetics associated with McDonald's consisted of its postmodern signification, and, in this case, claimed that that signification was distinct from beauty or elegance. I believe that within the context of the signification of these products there are components of beauty and elegance that McDonald's workers and their supervisor attempt to achieve.

> and chopped spinach with cream sauce [at home], the trout is completely boned and stuffed inside it, and it's wrapped in puff pastry and baked and served with a beurre blanc or vin blanc sauce. It's really a beautiful dish 'cause you make little puff pastry fish, and I'd like to try that.
>
> Author: Have you spoken to [the chef] about that?
> Cook: No, I'm just waiting. (Interview, La Pomme de Terre)

The chef and sous chef, more experienced, are expected routinely to create dishes. Even when cooks are permitted to innovate, they usually check with their supervisors. Once I asked the head day cook at La Pomme de Terre about dishes that had failed. He indicated that this does not often happen because: "We play it pretty safe. If it's outlandish, we ask [the head chef]" (field notes, La Pomme de Terre). Inexperienced cooks, with less autonomy, must acquiesce to the dictates of those higher up:

> Bruce, a regular evening cook at the Owl's Nest, complains about how the head chef makes him cook asparagus: "I hate lemon on asparagus. . . . It's all right, but it's not my taste, but it's what Paul likes He puts a whole rind in [while cooking], and it falls apart and goes all over the asparagus." (Field notes, Owl's Nest)

Status and role direct the locus of aesthetic decision making in the kitchen. The objects of production become the basis for reinforcing authority relations. Occupational segmentation means that not all have equal opportunity to participate in making these choices. Although there is a possibility within the kitchen for negotiation or at least a questioning of higher authority, an obdurate power structure determines what is served.

Occupational Task

Within any job, tasks vary. Some tasks involve a greater consciousness of the sensory dimension of production than others. Painting the background of a portrait is less aesthetically demanding than painting the figure, even though some aesthetic sensibility adheres to both. Some surgery is routine, while other surgery requires a light touch. Buffets and work on platters often involve close attention to appearances, while at other times aesthetic choices set by others affect the work. One hotel cook distinguished between creativity involved in working the line (preparing food to order) and planning a banquet plate: "A line has no creativity to it at all. As far as working in the back, I think you must have creativity because you always have to think up something creative to garnish up your plate with or to make your food look nice" (interview, Blakemore Hotel). Within an occupational routine, tasks differ in the

attention given them; this is, in part, a function of how much control the cook has over the contents of the plate or platter and how it is arranged.

Many cooks have interchangeable jobs. They switch tasks depending on immediate needs; they are not specialists.[19] Yet, specialty areas exist: notably that of pastry work, where the visual appeal of the dish is critical. The pastry chef at La Pomme de Terre defined the difference between cooking and pastry work as the difference between two art worlds: "I think people that get into pastry really heavily and do a lot of fancy decorating, and that's an art like painting. Whereas cooking has more artistic talent in preparing it to the proper degree of doneness and, plus, its arrangement on a plate, so it's a little bit more like photography" (interview, La Pomme de Terre). The great 19th-century French chef Carême linked pastry and architecture as one of the five fine arts (Revel 1982, p. 68). Pastry work, with a larger amount of unpressured time for preparation and planning, permits more thoughtful attention to aesthetic concerns than does "line" cooking.

The concern with the sensory qualities of products is a variable characteristic of occupations. While aesthetics is always present, its form and prominence differs. The status and market niche of an organization, the stage of one's career, and the particular task that must be completed, each influences how workers address their aesthetic concerns. These choices cannot be reduced to organizational demands, but they are channeled and specified by organizational and occupational characteristics.

BEYOND THE KITCHEN

A concern with the sensory qualities of products and production applies to all work life, not just restaurants. Much of what we mean by quality has this sensory (aesthetic) dimension; we suggest that the object (or performance) transcends functional requirements. Even when we are not self-conscious about stylistic components, we still care about what we produce and how we produce it. In this, all work has the components of artistic endeavors. House painters, portrait painters, and abstract expressionists have an aesthetic sensibility—a sense that the sensory characteristics of their products matter and that, ideally, the basis for evaluation should be determined by the group.

This does not deny the power of constraints. Structural constraints (production dynamics) mute an aesthetic centrality. The constraints may

[19] Obviously the status of the restaurant makes a difference in the fluidity of the division of labor, but in all these restaurants the head chef participates in doing routine work when needed.

derive from one's position, from one's clients, from the workplace dynamics, or from the organization's resource base. Art is like work and work is like art.

Examining aesthetic choices and their constraints expands the production of culture model. Production decisions are socially organized, but they are not *merely* a function of this organization. We require a sociology of work that treats aesthetic choices and decisions about quality as partially autonomous from production. I focused on the doings of professional cooks, but this analysis should be generalized to other occupations, even though details differ. Most occupations must confront the central demands of client control, organizational efficiency, resource management, and segmentation.

Client Demands

All practitioners realize they labor for those outside the occupation (Hughes 1971, p. 321). Even though clients rarely make explicit demands of the workers, the occasional complaint and the typification of the client constrains action. Lawyers (as well as their clients) are judged by juries; law clerks attempt to write beautiful briefs, barely read by put-upon judges (Riesman 1951). Dental patients care little about the dentists' standards for elegant fillings, as long as they do not feel pain and think they look good. Jazz musicians must put up with the frustrating ignorance of their audiences and shape their notes accordingly (Becker 1963, pp. 91–95). Ministers realize that God is not the only one judging their sermons (Kleinman 1984). In these cases, explicit demands are not made of the workers, but the messages filter through. After production is complete, evaluation begins, and the existence of audiences with different or ambiguous standards constrains activity. For some occupations clients continually judge subjects in which the worker has a greater expertise (e.g., cooking, hair styling, selling dresses) and this is seen in whether they return; for others the client is unconcerned or ignorant about the aesthetics of the work, provided the instrumental outcome and cost are satisfactory (e.g., plumbing, surgery).

Clients enforce their judgments when they consider the sensory appeal of the product or performance, and use that as a basis for further patronage. This is particularly evident in cases, such as food preparation, where clients receive quick and complete information in the form of the dish, as opposed to other production—such as auto repair—that is judged many miles down the road. When aesthetic choices "matter" to the clients, workers' decisions must address their taste; when clients do not care, these decisions are fettered by costs and efficiency. One of the

crucial goals of "professionalization" is to ensure that the primary source of evaluation for an occupation is internal, rather than external, and that clients accept this.[20]

Organizational Efficiency

The conditions of work, particularly temporal components, determine how much and what kind of things can be produced. Workers on an assembly line know that the line keeps moving. One has a limited time to do it right. Doing it right may be sacrificed to doing it. Writers have a cynical rule: "Don't get it right; get it writ." Court dates and judges' limits on closing statements pose a bar for attorneys. Patients can stand only so much anesthesia and parishioners plan Sunday dinners.

Some nuance of the task may be sacrificed because of the lack of patience of clients or because of the constraints on labor costs. The clock is a stern master, although the real master stands behind the clock. Workers in many venues negotiate to extend the time for completing work. While differences exist among occupations and segments of occupations, temporality has both a phenomenological and obdurate reality (Fine 1990).

Resource Management

The cost of materials sets a membrane around production. Ingredients, tools, and environments determine what can be done. The furniture upholsterer is at the mercy of the fabrics, the hairdresser at the mercy of the dyes, the sculptor depends on the quality of the marble, and the drill-press operator is limited by the machine. The quality of these resources is often out of the hands of the worker; it is decided upon by others with their own set of goals. All work is set within a market. The fit between resources and organizational environment places an obdurate brake on aesthetic choices.

Occupational Segmentation

Although all occupations must deal with the challenges posed by the constraints described above, differences within occupations also affect the doing of work. What you pay contributes to what you get. Hospitals, repair shops, architectural firms, and universities differ in the style and the competence of what is produced. In offices and organizations some

---

[20] We have no equivalent for the term "professionalization" for crafts, but the occupational autonomy among craftworkers points to the same issue.

are newer to the job, some have more autonomy, some care more, and some have positions that demand more conscious care: house painters are more conscious of the sensory effects of their work than industrial painters, surgeons more than anesthesiologists, or jockeys more than stablehands. Occupations are socially segmented, and different segments rely on different standards of judgment.

While aesthetic choice is a regular part of the doing of work, it is a variable, not an absolute. Both the centrality and amplitude of aesthetic interest must be recognized. These concerns coexist with keeping one's job, having the job be tolerably easy, and gaining self-esteem and material rewards. While each occupation has areas in which expressive choices are relevant, few totally lack such concerns. In contrast, no occupation is so devoted to the pursuit of form over function that social constraints do not exist. Factory work has a creative component (e.g., Bell 1984), just as artistry shows constraints of market and control systems that affect the doing of this ostensibly "purely" creative work.

THE CULTURE OF PRODUCTION

Management and labor are in firm agreement that work quality is crucial. Aesthetic production should be consistent with organizational goals, not subversive of them. Yet, the intersection of the expression of quality may produce friction. Workers wish that they had more time (implying they need more co-workers) and more resources, so they can produce in an unhurried fashion. Management is likely to emphasize greater efficiency. Good work is profitable to a point, and this point is connected to market niche and price elasticity. Management has the direct problem of profitability, whereas for workers, profitability is only an indirect concern. As a result, value consensus may devolve into conflict or frustration in actual practice.

To the extent that workers have and can maintain a craft orientation, they can extend their zone of discretion in production decisions. To the extent that they are connected into a bureaucratic organization, management makes the choices, solidified into rules and procedures, that workers carry out. A strain exists between the craft organization of work, which vests authority with the members of the occupation, and the bureaucratic organization of work, where decisions are a result of authority hierarchies and formal procedures. Occupations in which each object is uniquely prepared reinforce the craft orientation; jobs that emphasize consistency and efficiency tend to be found in bureaucratic organizations (Stinchcombe 1959). Even in the latter arenas, management may tolerate, even encourage, some worker discretion if, although it does not maximize profits, it reduces labor discontent and allows for a predictable flow of

production (Burawoy 1979). The role of discretion is indicated by the willingness of management to permit cooks to take extended breaks, shift positions, and choose which dishes to recook. The effects of this light hand are seen in cooks' willingness to work overtime (or to come in early), fill in for absent others, and make special dishes for important customers—each beyond the limits of formal job requirements. Furthermore, when worker aesthetics are congruent with that of management, some flexibility on material and labor costs may be tolerated, and passed on to the customer as the inevitable expense of quality.

Producers, consumers, and managers all value good work within imperatives of monetary or psychic costs. When the system is working, each is willing to accede at critical points. The challenge for management, especially evident at La Pomme de Terre, the most explicitly artistic of the sites, is to have workers accept management's vision of material constraints as a given, and to work within those constraints. Since there is a trade-off in quality and cost, mediated by customer evaluation, the choices are not objective. Organizational success in expressive production involves a moving dynamic: to be good enough *and* cheap enough that one's targeted customers will return and recruit others.

This analysis suggests the importance of transcending the commonplace that art is like all work, but it also shows where, when, and how aesthetic autonomy and social control interpenetrate and how they are negotiated. Under which circumstances do workers have concerns about the sensory quality of their products and services and when are they permitted control over this quality? The answer is shaped by the situated reality of workplace negotiation and by the reality and the typification of the market.

The sociological treatment of the expressive side of production remains largely unmapped. A single case can only provide outlines for others to fill. Specifically the causes of particular aesthetic choices have been ignored. How do workers derive an understanding of what is right and valued? What dimensions—instrumental and expressive—determine quality of production? How is cultural capital generated in work? Under what circumstances is elegant simplicity valued? When is self-conscious creation of the beautiful crucial? Issues of the aesthetics of performance and the aesthetics of products need to be differentiated. Finally, comparative research on numerous occupations avoids a haziness of the description of aesthetic choices.

The emphasis on and expression of aesthetic choices depends on the work environment, the standing of the worker, and the particular work task. Workers' orientations to the expressive side of production are grounded in the core sociological concepts of contention, autonomy, and community; management's limitations are equally sociological, based on

demands for control and efficiency deriving from instrumental requirements. Work is a minuet between (expressive) form and (instrumental) function. In this dance, as in others, he who pays the piper ultimately calls the tune.

REFERENCES

Aldrich, Virgil. 1966. *Philosophy of Art*. Englewood Cliffs, N.J.: Prentice-Hall.
Arian, Edward. 1971. *Bach, Beethoven, and Bureaucracy*. University· University of Alabama Press
Beardsley, M. C. 1958. *Aesthetics: Problems in the Philosophy of Criticism*. New York. Harcourt Brace.
Becker, Howard S 1963. *Outsiders*. New York: Free Press.
———. 1974 "Art as Collective Action." *American Sociological Review* 39:767–76.
——— 1982. *Art Worlds* Berkeley: University of California Press.
Bell, Michael J 1984. "Making Art Work." *Western Folklore* 43:211–21.
Blau, Judith. 1984 *Architects and Firms*. Cambridge, Mass.: MIT Press.
Bourdieu, Pierre. 1984. *Distinction*. Cambridge, Mass.: Harvard University Press.
Bucher, Rue. 1962. "Pathology: A Study of Social Movements within a Profession." *Social Problems* 10:40–51.
Burawoy, Michael. 1979. *Manufacturing Consent*. Chicago. University of Chicago Press.
Clark, Priscilla. 1975. "Thoughts for Food I: French Cuisine and French Culture." *French Review* 49:32–41.
Coser, Lewis, Charles Kadushin, and Walter Powell. 1982. *Books*. New York: Basic.
Csikszentmihalyi, Mihalyi. 1975. *Beyond Boredom and Anxiety*. San Francisco· Jossey-Bass.
Danto, Arthur. 1964. "The Artworld." *Journal of Philosophy* 61.571–84
———. 1981. *The Transfiguration of the Commonplace*. Cambridge, Mass.: Harvard University Press
Dickie, George. 1974. *Art and the Aesthetic* Ithaca, N.Y.: Cornell University Press.
Diffey, T. J. 1984. "The Sociological Challenge to Aesthetics." *British Journal of Aesthetics* 24.168–71.
Dimaggio, Paul. 1977. "Market Structure, the Creative Process, and Popular Culture." *Journal of Popular Culture* 11.436–52.
Faulkner, Robert. 1971. *Hollywood Studio Musicians* Chicago: Aldine.
Fine, Gary Alan. 1979 "Small Groups and Cultural Creation The Idioculture of Little League Baseball Teams " *American Sociological Review* 44 733–45.
——— 1982. "Multiple Socialization: The Rhetorics of Professional Cooking." Paper presented to the American Educational Research Association, New York.
———. 1985. "Occupational Aesthetics: How Trade School Students Learn to Cook." *Urban Life* 14:3–32.
———. 1987. "Working Cooks: The Dynamics of Professional Kitchens." Pp. 141–58 in *Current Research on Occupations and Professions,* edited by Helena Z Lopata. Greenwich, Conn: JAI.
——— 1989. "Wittgenstein in the Kitchen: The Creation of Meanings of Talk." Paper presented to the Midwest Sociological Society, St. Louis.
———. 1990. "Organizational Time· The Temporal Experience of Restaurant Kitchens." *Social Forces* 69·95–114.
Finkelstein, Joanne. 1989. *Dining Out: A Sociology of Manners* New York: New York University Press.

Freeman, John, and Michael Hannan. 1983. "Niche Width and the Dynamics of Organizational Populations." *American Journal of Sociology* 88:1116–45.

Garfinkel, Harold. 1967. *Studies in Ethnomethodology.* Englewood Cliffs, N.J.· Prentice-Hall

Griswold, Wendy. 1986. *Renaissance Revivals.* Chicago: University of Chicago Press

Gross, Edward. 1958. *Work and Society.* New York. Crowell

Hall, Richard H. 1975. *Occupations and the Social Structure,* 2d ed. Englewood Cliffs, N.J.: Prentice-Hall.

Hincks, Tony 1984 "Aesthetics and the Sociology of Art: A Critical Commentary on the Writings of Janet Wolff." *British Journal of Aesthetics* 24:341–54.

Hirsch, Paul M. 1972. "Processing Fads and Fashions. An Organization-Set Analysis of Cultural Industry Systems." *American Journal of Sociology* 77:639–59.

Hughes, Everett. 1971. *The Sociological Eye.* Chicago: Aldine.

Kant, Immanuel. 1952 *Critique of Aesthetic Judgment.* Oxford. Oxford University Press

Kimball, Christopher. 1985. "The *Cook's* Interview· Jeremiah Tower." *The Cook's Magazine* (Jan./Feb)·18–19

Kleinman, Sherryl. 1984. *Equals before God* Chicago. University of Chicago Press.

Lauer, Robert H 1981. *Temporal Man* New York: Praeger.

Moir, H. C. 1936. "Some Observations on the Appreciation of Flavour in Foodstuffs." *Chemistry and Industry* 55:145–48.

Orwell, George. 1933. *Down and Out in Paris and London.* New York: Harcourt Brace

Pangborn, R. M. 1960. "Taste Interrelationships." *Food Research* 25:245–56.

Peterson, Richard A. 1979. "Revitalizing the Culture Concept." *Annual Review of Sociology* 5:137–66.

Revel, Jean-Francois. 1982. *Culture and Cuisine.* Garden City, N.J.. Doubleday.

Riesman, David. 1951. "Toward an Anthropological Science of Law and the Legal Profession." *American Journal of Sociology* 57:121–35.

Sclafani, Richard J. 1979. "Artworks, Art Theory, and the Artworld." *Theoria* 39.18–34.

Shepard, Anne. 1987. *Aesthetics* Oxford: Oxford University Press.

Shelton, Allen. 1990. "A Theater for Eating, Looking, and Thinking: The Restaurant as Symbolic Space." *Sociological Spectrum* 10 507–26.

Stewart, Doug. 1988. "For a Portraitist, Making Faces Is a Hard Day's Fight." *Smithsonian* (July): 43–50.

Stolnitz, Jerome. 1960. *Aesthetics and Philosophy of Art Criticism.* Boston: Houghton Mifflin.

Stinchcombe, Arthur. 1959. "Bureaucratic and Craft Administration of Production: A Comparative Study." *Administrative Science Quarterly* 4:168–87.

Strauss, Anselm. 1978. *Negotiations.* San Francisco: Jossey-Bass.

Walker, Charles R., and Robert H Guest. 1952. *The Man on the Assembly Line.* Cambridge, Mass.: Harvard University Press.

Whyte, William Foote. 1948. *Human Relations in the Restaurant Industry.* New York: McGraw Hill.

Wolff, Janet. 1983. *Aesthetics and the Sociology of Art.* Boston: Allen & Unwin.

Woodward, Joan. 1965. *Industrial Organization: Theory and Practice.* London: Oxford University Press.

Repair after Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation[1]

Emanuel A. Schegloff
University of California, Los Angeles

Organizational features of ordinary conversation and other talk-in-interaction provide for the routine display of participants' understandings of one anothers' conduct and of the field of action, thereby building in a routine grounding for intersubjectivity. This same organization provides interactants the resources for recognizing breakdowns of intersubjectivity and for repairing them. This article sets the concern with intersubjectivity in theoretical context, sketches the organization by which it is grounded and defended in ordinary interaction, describes the practices by which trouble in understanding is dealt with, and illustrates what happens when this organization fails to function. Some consequences for contemporary theory and inquiry are suggested.

CONTEXTS FOR INTERSUBJECTIVITY

Theoretical Context

Virtually all social theory has presupposed (and some has explicitly recognized) that underlying the very conception of a social unit—whether group, class, or society—and the very conception of orderly social process is some common grasp by the unit's members of their common situation of action and of the import of ordinary conduct within it. Although this grasp might variously be seen as a consensually shared outlook or as a

---

view imposed by some set of persons and interests, some co-conception or coorientation to the world has seemed inescapable.

In focusing on the relationship between repair after next turn and "intersubjectivity," this article underscores a connection between talk-in-interaction as a primordial site of sociality on the one hand and, on the other hand, one of the (largely presupposed) preconditions for, and achievements of, organized social life. On the whole, the problem of intersubjectivity, derivative from the individualism and atomism of Western, Judeo-Christian culture and aggravated by the privatization of mind precipitated by Cartesian and post-Cartesian philosophy, has been addressed in Western theorizing largely by philosophy. For sociology, though recognized in principle, the issues posed by the problematic character of intersubjectivity have been largely ignored in practice.

However, the problematics of intersubjectivity are anterior to most of the problems that sociological and social theory have treated as primary and fundamental, such as the so-called Hobbesian problem of order or the underlying engines of large-scale social change. Most simply put, without systematic provision for a world known and held in common by some collectivity of persons, one has not a misunderstood world, but no conjoint reality at all. That is, the problem of intersubjectivity (or cognitive order) is theoretically anterior to whatever formulations of problems of order or conflict are part of the tradition of social theory. Absent intersubjectivity, the terms of any social theory—whether they refer to interests or values, persons or roles, authority or power—by definition cannot name anything oriented to or effective with any regularity or commonality, for there could not be any common recognition of them.

A very general type of solution has typically been invoked in social theory to provide for intersubjectivity. This "solution" turns on "common culture" as the social resource by which the individual's grasp of reality is mediated. Intersubjectivity has been understood to be under-girded first by the transmission of culture through socialization (primarily through primary socialization in childhood), together with the segregation of social units (subgroups or subcultures) whose cultural resources diverged substantially. Where these devices for ensuring intersubjectivity failed, characteristic forms of "disorganization" might surface—in mental illness and other forms of deviance, as well as in more or less overt conflict between groups with divergent "understandings of the world."

In sociology, the main theoretical lineage that has carried this stance has been that running from Durkheim (esp. *The Elementary Forms of the Religious Life*, 1915) to Parsons. Although rarely made explicit, the fundamental conception of such a common culture or shared understanding of the world (as was made clear by Garfinkel [1967] in the course of challenging it) was of more or less identical contents of separate minds.

Members of a culture (or subculture), upon being furnished the same fundamental conception of time, space, and causality (to recall Durkheim's concerns in coming to terms with Kant's fundamental categories of mind) as well as the remaining conceptual grid by which the surrounding real world was categorized and rendered orderly, would come to apperceive the world in roughly congruent ways. Thus was the starting point of *individual* minds encountering an *external* world (including other individuals) rendered compatible with a consensual or intersubjective grasp of reality, both physical and social (cf. Heritage, 1984*a*, chap. 2).

Not that this was addressed directly as a problem in sociological theory. Although Parsons (1937, chap. 11) worried somewhat about the problems of epistemic relativism that might result from Durkheim's effort to "socialize" Kant's fundamental (i.e., irreducible) categories of mind, he escaped seriously grappling with this class of problems largely by dismissing this aspect of Durkheim's work.

The passage of a more serious and sustained concern with intersubjectivity from philosophy to sociology was undertaken, with variable success, for pragmatism most centrally by George Herbert Mead (1938, 1962; and see Joas 1985) and for phenomenology by Alfred Schutz (e.g., [1932] 1967; or 1962, pp. 3–47). For example, Schutz begins by insisting that commonsense thinking is situated "from the outset [in] an intersubjective world of culture" (1962, p. 10), and he goes on to focus specifically on the observation that the "world is not my private world but an intersubjective one and that, therefore, my knowledge of it is not my private affair but from the outset intersubjective and socialized" (p. 11). Explicit recognition of such a property allows explicit attention to how socialized knowledge has its intersubjectivity provided for, for example, by the feature of commonsense knowledge that Schutz speaks of as "the reciprocity of perspectives" (p. 14).

In contemporary sociology, it was Garfinkel who most forcefully brought to attention the centrality to social theorizing of a world-known-in-common and of commonsense knowledge, and it was he who showed the inadequacy of the largely tacit conceptions of it that underlie most contemporary theorizing. In a series of studies (collected in Garfinkel [1967]) prompted in the first instance by confronting the work of Talcott Parsons with the most sociologically relevant strands of phenomenology (see Heritage [1984*a*] for a lucid account of the theoretical lineages and interactions here), Garfinkel asked what exactly might be seriously intended by such notions as "common" or "shared" knowledge. In the days when computers were still UNIVACs, he showed as untenable that notion of "common" or "shared" that was more or less equal to the claim that separate memory drums had identical contents. When even

the sense of ordinary words and very simple sentences could be shown not to engender identical explications when presented to different persons, when those explications had themselves to be reconciled to provide them a "sense of equivalence," and when *those* reconciliations in turn required such reconciliation, and so on, the notion of "common culture" or "shared knowledge" as composed of same substantive components—whether norms or propositions—being "held" by different persons became increasingly difficult to defend.

Instead, what seemed programmatically promising was a *procedural* sense of "common" or "shared," a set of practices by which actions and stances could be composed in a fashion which displayed grounding in, and orientation to, "knowledge held in common"—knowledge that might thereby be reconfirmed, modified, expanded, and so on. Garfinkel's term "ethnomethodology"—with its explicit preoccupation with the procedures by which commonsense knowledge is acquired, confirmed, revised, and so on—can be partially understood by reference to this matrix of concerns. As Garfinkel wrote in a chapter entitled "What Is Ethnomethodology?" (1967, p. 30), " 'Shared agreement' refers to various social methods for accomplishing the member's recognition that something was said-according-to-a-rule and not the demonstrable matching of substantive matters. The appropriate image of a common understanding is therefore an operation rather than a common intersection of overlapping sets."

Garfinkel drew heavily on Schutz (cf., e.g., in "Studies of the Routine Grounds of Everyday Activities" [Garfinkel 1967]) in describing the operational character of commonsense knowledge. The intersubjective character of commonsense knowledge is undergirded by a collection of features provided by a variety of interpretive procedures constituting the so-called attitude of daily life. Applicable to any particulars of social settings, the elements of the solution to the problem of intersubjectivity are principled, interpretive operations that constitute the "mundane" grasp of the world (Pollner 1987).

This general stance (by which I mean not only Garfinkel's, but that of others for whom some version of interpretive procedures is central, such as Cicourel [1972, 1974], Gumperz [1982], or Pollner [1987]) appears designed to disavow access to any determinate structure or character of a real world of input—whether from the physical world or from the conduct of other social actors—and to focus on interpretive procedures as, in effect, the sole locus and source of interpreted order.

An alternative, or supplementary, stance might take the conduct of other social actors as not, in effect, random or inaccessible to affirmative inquiry, but, rather, together with interpretive procedures, coshaping an appreciated grasp of the world. Such a view would allow for the interven-

tion by the accountable authors of conduct in what would come to be stabilized as the effective understanding of that conduct. Intersubjectivity would not, then, be merely convergence between multiple interpreters of the world (whether understood substantively or procedurally) but potentially convergence between the "doers" of an action or bit of conduct and its recipients, as coproducers of an increment of interactional and social reality.

In the context of such a stance, intersubjectivity is not a matter of a generalized intersection of beliefs or knowledge, or procedures for generating them. Nor does it arise as "a problem of intersubjectivity." Rather, particular aspects of particular bits of conduct that compose the warp and weft of ordinary social life provide occasions and resources for understanding, which can also issue in problematic understandings. And it is this situating of intersubjectivity that will be of interest here.[2]

The achievement and maintenance of this sort of intersubjectivity is not treated in a theoretically satisfactory manner by invoking socialization as a mechanism, for intersubjectivity is achieved for a virtually inexhaustable range of types of events always contextually specified, for which no "distal" or "remote" socialization could provide. The solution surely is provided for by a resource that is itself built into the fabric of social conduct, into the procedural infrastructure of interaction. For the domain of conduct addressed in this article, this involves a self-righting mechanism built in as an integral part of the organization of talk-in-interaction—what has been termed the organization of repair. I try to show how the procedural basis for locating and dealing with breakdowns in intersubjectivity is woven into the very warp and weft of ordinary conversation and, by implication, possibly of any organized conduct.

Grounding in the Sequential Organization of Conversation

The sequential basis for the (largely unnoticed) confirmation of intersubjectivity, and for its occasional repair in "third position," can be very briefly sketched. In turns at talk in ordinary conversation, speakers ordinarily address themselves to prior talk, and, most commonly, to immediately preceding talk. In doing so, speakers reveal aspects of their under-

---

[2] See also Goffman (1983) for an examination of common understandings through the treatment of "presupposition," "deixis," and "background expectations" by linguists, sociolinguists, and sociologists. Goffman does remark in passing (p. 46) on the possibilities of *mis*understanding, but relates these more to the vulnerabilities of cryptic utterances between intimates than to the systematic contingencies of the repair of misunderstanding in general (my thanks to Candace West for reminding me of the related theme of Goffman's paper.)

standing of the prior talk to which their own speech is addressed (Sacks 1989, in press; Moerman and Sacks as presented in Moerman [1988]).

For example, in responding to an interlocutor's "How are you?" with "Fine," a speaker can manifest a variety of simple, but fundamental, understandings. In starting to talk, he or she can show an understanding that the prior speaker's turn was possibly finished. In producing an answer, this same speaker can display an understanding of the prior utterance as a question. By forming up the answer through a description of personal state (or "adjective of manner"), the speaker can show an understanding of the grammatical type of the question as a so-called WH—— question (i.e., of the type beginning with WH as in *who, where,* etc.) as compared to a yes/no question. By selecting this particular "value" of response (as compared, e.g., to "terrible," "wonderful," or an actual launching of a story) the speaker can reveal an understanding of the force or standing of such a question in the interaction in progress, in view of the type of interaction, the state of the relationship of the participants, and so on (see Sacks 1975; Jefferson 1980; Schegloff 1986).

Similarly, with the rejoinder, "That's good," the initial speaker can in turn display understandings that the "responder's" turn is possibly complete, that it was indeed a response to the question that had preceded it (and was not directed to some third party, e.g., or delivered as part of a mumbled internal dialogue), that it was designed to report a favorable "state" (hence the positive assessment in return), and to pass an opportunity to undertake some extended report (Sacks 1975; Jefferson 1980; Schegloff 1986).

And so on. Each next turn provides a locus for the display of many understandings by its speaker—understandings of what has immediately preceded (and of far less readily apparent features than those used here as illustrations; see Schegloff [1984, 1988b, pp. 118–25, 1990] for illustrative, more elaborate, discussions) or of what has occurred earlier or elsewhere that nonetheless figures in the turn's talk. The understandings are displayed en passant for the most part (although there is also a distinct type of utterance overtly designed to check its speaker's understanding of preceding talk), as by-products of bits of talk designed in the first instance to do some action such as agreeing, answering, assessing, responding, requesting, and so on.

Having registered the observation that, through their talk, speakers can display aspects of their understanding of prior talk, it remains to be noted that, in doing so, they can reveal understandings that the speakers of that prior talk find problematic—in other words, what they take to be *mis*understandings. There can be misunderstandings of what is being referred to—sometimes because a word or phrase or usage is accessible to alternative interpretation (as in the case of pronouns and other deictic

or indexical expressions), sometimes in the face of quite explicit and "readily hearable" references (see excerpt 18 below). There can be misunderstandings of the upshot of what a speaker is doing with a turn's talk—for example, misunderstandings on the serious/nonserious dimension, such as taking a joke seriously or vice versa, or taking as a complaint something claimed as otherwise intended (for these and other types of misunderstanding, see Schegloff [1987a]).

When such "problematic understandings" occur, and whatever their apparent "source," speakers of the "misunderstood" talk can undertake to "repair" the misunderstanding, and this can thus constitute "third position repair"—repair after an interlocutor's response (second position) has revealed trouble in understanding an earlier turn (the "repairable" in first position). The ordinary sequential organization of conversation thus provides for displays of mutual understanding and problems therein—one running basis for the cultivation and grounding of intersubjectivity (see App. A below).

Third position repair may be thought of as the last systematically provided opportunity to catch (among other troubles) such divergent understandings as embody breakdowns of intersubjectivity—that is, trouble in the socially shared grasp *of the talk and the other conduct in the interaction.*

THIRD POSITION REPAIR

Dealing with Trouble in Understanding

Participant attention to troubles in understanding is deployed in all the positions at which we find attention to other sorts of trouble in conversation.

Thus, for example, speakers of a turn may orient themselves to prospective problems in its being understood and may build into their talk—while the turn is still in progress and incomplete or, just after its possible completion, in what is called the "transition space" (Sacks, Schegloff, and Jefferson 1974, pp. 705–6)—resources addressed to such potential problems.

For example, in the following utterance Marcia, in explaining to her ex-husband why their son is flying rather than driving back home, appears to pick up the possible ambiguity of the phrase "ripped off," as between the literal tearing of a convertible car's soft top and the idiomatic usage for robbery.

Excerpt 1 (MDE, MTRAC, 3·1)
Marcia: . . Becuz the to.p was ripped off'v iz car which
           iz tihsay someb'dy helped th'mselfs.

"Which is to say somebody helped themselves" seems directed to "disambiguating" this possible source of misunderstanding. Efforts by the speaker of some talk to deal with problems of understanding thus can be initiated in, or just after, the very turn in which the talk engendering these troubles occurs, as is the case with other sorts of troubles that talk can give rise to.

But if a speaker does not address potential troubles of understanding, or if a recipient of the talk claims such problems whether or not they have been addressed by the speaker, efforts to deal with problems of understanding can be initiated by the *hearer*. Virtually all such efforts (see App. A) are initiated in one place—the turn *after* the turn in which the source of the trouble occurred. (Note that I am referring here to the *initiation* of efforts to deal with some trouble; carrying through such repair may extend past the turn in which the repair is initiated.)

Thus, in the episode already cited, Marcia's effort to explicate her usage of "ripped off" is itself not immune to problems of understanding (perhaps because of her ironic figure), and this problem is then addressed in next turn by the talk's recipient, who offers a candidate understanding, which Marcia then confirms.

Excerpt 2 (MDE, MTRAC, 3·1)
Marcia: . . . Becuz the to:p was ripped off'v iz car which
           iz tihsay someb'dy helped th'mselfs.
  Tony: Stolen.
           (0.4)
Marcia: Stolen. Right out in front of my house.

Given, first, the detailed capacity of speakers for designing their talk for the context and the moment in which it is being done, including its current-at-that-moment recipients (Goodwin 1979, 1981), given, second, the self-repair exercised by speakers on their own talk within and just after their turns, and given, third, the recipients' next-turn addressing of troubles that have nonetheless crept into the talk, the vast majority of understanding problems are dealt with virtually "immediately," as is the case with other types of troubles that can come up. Still, there is a recurrent set of circumstances, outlined in excerpt 3, that escape these processes.

What Is "Third Position Repair"?

Excerpt 3
A: Turn 1   (T1)
B: Turn 2   (T2)
A: ←

It happens that a speaker of a turn, $T1$, "releases" it as adequate, and its recipient finds in it no problem that warrants initiating repair in the next turn position. Accordingly, its recipient produces a next turn, $T2$, sequentially appropriate to his or her understanding of what the speaker of the prior turn was doing in $T1$ and reflecting his or her understanding of what the prior speaker may have been referring to by various referring terms in $T1$ (a process treated by Clark and Schaefer [1989] under the term "grounding"). And $T2$, built to be and understood as "responsive" to $T1$, thus regularly displays to the speaker of the prior turn the understanding that has been accorded it—an understanding that the speaker of $T1$ may treat as problematic.[3]

After such meant-to-be-sequentially-appropriate next turns, in what we can term third position (where the misunderstood talk's turn is first and the responsive next turn is second), the speaker of the problematically understood talk—the trouble source—can undertake to address the trouble by engaging in some operation on the source of the trouble, that is, the talk in $T1$. Several instances may serve to provide empirical displays of this otherwise abstract schema.

In excerpt 4, the press relations officer in a Civil Defense headquarters is asking the chief engineer for information to be distributed to the media (see App. B for transcription conventions).

Excerpt 4 (CDHQ, I, 52)

Annie:	Which one::s are closed, an' which ones are open.
Zebrach:	Most of 'em. This, this, ⌜this, this ((pointing))
→ Annie·	⌞I 'on't mean on the
→	shelters, I mean on the roads.
Zebrach:	Oh!
	(8.0)
Zebrach:	Closed, those're the ones you wanna know about,
Annie:	Mm⌜hm
Zebrach:	⌞Broadway . . .

In excerpt 5, the therapist in a group therapy session for teenagers offers an observation on the dynamics of the discussion.

Excerpt 5 (GTS, I, 37)

Dan:	Well that's a little different from last week.
Louise:	heh heh heh Yeah. We were in hysterics last week.
→ Dan	No, I mean Al.
Louise:	Oh. He . . .

In excerpt 6, while visiting the city in which her old friend Alice now

---

[3] An account of some of the recurrent sources and types of "misunderstanding" addressed in these contexts may be found in Schegloff (1987a).

lives, Belle has called to say hello, and the talk turns to the prospect of
their getting together.

Excerpt 6 (DA, 2)

	Alice.	Well I'd like tuh see you very much.
	Belle:	Yes. ⌜Uh
	Alice:	⌞I really would. We c'd have a bite,
		en ⌜(ta::lk),
	Belle:	⌞Yeh.
	Belle:	Weh- No! No, don't prepare any⌜thing.
	Alice:	⌞And uh-
	Alice:	I'm not gunnuh prepare, we'll juz' whatever
		it'll ⌜be, we'll
→	Belle:	⌞NO!
→	Belle:	I don' mean that. I min- because uh, she en I'll
→		prob'ly uh be spending the day togethuh, so uh:::
→		we'll go out tuh lunch, or something like that.
→		·hh So I mean if you·: uh have a cuppa coffee or
→		something, I mean ⌜that uh that'll be fine. But=
	Alice:	⌞Yeah
→	Belle:	=⌜uh- othuh th'n that don't ⌜uh- don't bothuh=
	Alice:	⌞Fine. ⌞( )
→	Belle:	=with anything else.

In each of these cases, the arrows point to what I am terming "third
position repairs."

This is the major sequential context, then, for what we may relevantly
term "repair after next turn."[4] In what follows, I want to focus attention
on this analytically specified, but continually shifting, sequential position
in conversation and the character of the repair that is initiated there.
Repair in this position supplies, and is dedicated largely to, what I have
already referred to as "the last structurally provided defense of intersub-
jectivity in conversation."

Composition of Third Position Repair

Third position repairs have a highly recurrent form and are constructed
from four main types of components, some of which are themselves real-
ized by a very few types of lexical tokens.[5] I will initially refer to the

[4] I call repairs undertaken in a next turn by the same speaker, and which merely
happen to be after a next turn but are not in that position in an organizationally
criterial way, "third *turn* repairs" in contrast to the third *position* repairs examined
here (Schegloff, in press *b*).

[5] In discussing the canonical form taken by third position repairs, some wonder
whether I have ignored third position repairs that did not happen to be implemented
in the canonical format. Excerpt 20, presented below as an example of fourth position
repair, struck some as an instance of third position repair in deviant format. Else-

four components unimaginatively as *A*, *B*, *C*, and *D*, complementing these with more descriptive characterizations as they are introduced in turn. As is implied by using a set of ordered terms even for the preliminary references, the components that are employed (and not every component is employed in every third position repair) have a canonical ordering. As I will try to show, this ordering is the product of determinate practices of constructing these repair turns. Limitations of space, here as elsewhere in this presentation, constrain all the elements of the account to a minimum.

The *A* component of third-position-repair turns serves to initiate the repair. Most commonly it takes the form of "no," singly or in multiples ("no, no"; "no, no no"), or in combination with "oh" ("oh no"), which also occasionally stands alone ("oh").[6] Since excerpts 5 and 6 above, and most excerpts to be cited below, provide exemplars of these turn-initial particles, I will forgo separate displays of this "repair initiating component" here.

The *B* component is the one most likely to be absent from a third position repair. In fact, this component—an agreement/acceptance component—occurs virtually exclusively when the *T*2 (the next turn) has treated the *T*1 (the trouble-source turn) as a complaint, and the speaker of *T*1 has responded with some sequentially implicated response type, such as an apology, an excuse, and so on. Then the *B* component serves to agree with or accept the "response to the complaint," even though its speaker is about to go on to deny that his or her prior turn was doing complaining in the first instance.

Thus, in excerpt number 7, Agnes and Portia, sisters in their fifties, have several times missed getting together, and the telephone conversation from which this excerpt is taken began with a comment by Portia about another such failure. In Portia's first turn in this excerpt she is

---

where, I examine several instances of third position repair that are indeed differently formatted (see Schegloff 1991*b*).

[6] "Well" is also used as an *A* component, but in all the cases I have seen, the turn immediately proceeds to the *D* component—the repair proper. Such turns then take the form "Well I mean . ." as in excerpt 12 below. Although third position repairs may initially appear to be disagreements with the prior turn, and no may appear to signal such disagreement, it should be noted, first, that in the remainder of these turns the speakers operate *on their own prior talk,* not on that of the other, and, second, that prior research (Sacks 1987; Pomerantz 1984) has shown that disagreements do not ordinarily have disagreement tokens directly after a prior turn and at the start of the disagreeing turn. The first component of third position repair, with "no" as its turn initial particle, is best understood as initiating repair, rather than as betokening disagreement. Indeed, it is a way of beginning to constitute the turn *as* a third position repair. Constituting it as a disagreement is done in other ways (departures from this claim are discussed below)

apparently initiating the close of the conversation (see Schegloff and Sacks 1973; Button 1987*a*), but this is taken by Agnes as a reprise of the complaint about not getting together (Schegloff 1987*a*):

Excerpt 7 (NB)

```
    Agnes·    I love it.
       .      (0.2)
    Portia:   Well, honey? I'll pob'ly see yuh one a' these
              day:s.
    Agnes:    Oh.: God yeah,
    Portia.   ⌈Uhh huh!‾
    Agnes:    ⌊We-
    Agnes·    B't I c- I jis' ⌈couldn' git down ⌈there.
→   Portia:                   ⌊Oh-              ⌊Oh I know.
              I'm not askin ⌈yuh tuh ⌈come dow-‾
    Agnes:                  ⌊Jesus. ⌊I mean I jis'- I didn' have
              five minutes yesterday.
```

At the arrowed turn, the repair-initiating component ("Oh") is followed by an acceptance of Agnes's excuse for not visiting, the *B* component.

In 8, Bonnie has called Jim, her on-again off-again boyfriend, with a last-minute invitation to a New Year's Eve party, and Jim has declined to give a firm commitment, saying that he has already been invited to several other parties. Then:

Excerpt 8 (NYI, 6–7)

```
    Bonnie.   Because I'm not even sure if we're goin' to have
              it yet because a buncha people say ⌈maybe, maybe,
    Jim:                                         ⌊Yeah
    Bonnie.   't's buggin me.
              (1 5)‾
    Jim·      Oh uhh hh I'm sorry, Ihh
→   Bonnie·   No, that's okay, I mean y'know I can understand
              because- this was just a late idea that me and
              Barb had.
```

Again, at the arrowed turn a repair-initiation component ("No") is followed by an agreement/acceptance component ("That's okay") that accepts an apology responding to an apparent complaint.

The third, or *C*, component of third position repairs may be termed "the rejection component." With it, the speaker overtly rejects the understanding that prior turn reveals *its* speaker to have accorded the trouble-source turn. Two of the three main formats employed in this rejection component specify "by name" just what the repairer understands the misunderstanding to have been. They are used for the two most common types of misunderstanding (Schegloff 1987*a*): problematic reference and problematic sequential implicativeness (or what action a speaker has meant to be doing with the turn). For the former, the recur-

rently used form is "I don't mean X," as in excerpt 4 above and many other instances. For the latter, the form is "I'm not Xing," where X is the name of some action, as in excerpt 7 above ("I'm not askin you to come down") or in 9 below. Here, Dan—the therapist in a group therapy session with teenagers—has offered a characterization of one of them (Al), which is understood by another (Roger) as a critique or complaint.[7] Roger responds to the complaint against Al by asserting solidarity with his "buddy."

Excerpt 9 (GTS)

Dan:	. . See Al tends, it seems, to pull in one or two individuals on his side (there). This is part of his power drive, see. He's gotta pull in, he can't quite do it on his own. Yet.
Al:	W'l-
Roger:	Well so do I.
→ Dan:	Yeah. ⌈I'm not criticizing, I mean we'll just uh =
Roger:	⌊Oh you wanna talk about him.
Dan.	= look, let's just talk.
Roger:	Alright.

At the arrowed turn, after an agreement/acceptance component in which Dan agrees with Roger's response to the perceived complaint, Dan specifically rejects that displayed understanding of his prior turn ("I'm not criticizing").

It is striking that misunderstandings are both orderly and accessible to the speaker of what has been misunderstood, who might well be thought to be so committed to the design and so-called intent of the earlier turn as to be disabled from appreciating that (or how) it could be otherwise understood.[8] In the formats described so far, the product of this understanding of the misunderstanding is itself displayed by name. On other occasions, however, although the misunderstanding is overtly rejected, it is not named but is referred to by a pronoun—"I don't mean *that*" or "*That*'s not what I mean," as in excerpt 6 above. These do not appear

---

[7] Caution is necessary when drawing on material from settings that can be regarded as distinctive speech exchange systems, as therapy sessions often are. Such caution is especially relevant when modified turn-taking organizations may be involved (see n. 22 below). For a discussion of the serious analytic problems in formulating participants in the manner used in the text at this point, see Schegloff (1991a, in press a). Use of such characterizations introduces a measure of analytic informality into the present treatment, but does not appear to compromise the main issue being addressed by this article.

[8] Indeed, on occasion the misunderstanding is not accessible to the misunderstood speaker, with consequences that cannot be taken up here. See "A Breakdown of Intersubjectivity" below.

to reflect a failure to grasp the misunderstanding, only some potential trouble in formulating it appropriately.[9]

The *D* component might be termed "the repair proper." It is the component most likely to be present in any turn that is the locus of third position repair (although it too can be omitted, as in 7 above). This is the component in which the speaker carries out some operation or operations on a prior turn, so as to address the problematic understanding of it revealed by an interlocutor's response.

One form that the repair proper takes is a repeat of the trouble-source turn, produced in a manner that displays, or does, "saying it more clearly." In excerpt 10, James and Vic—two apartment house janitors in the Bronx—are discussing a broken pane of glass in a door in James's building, when he (James) comes across his income tax refund while opening his mail. (Mike is an employee in the used furniture shop where the conversation is taking place.)[10]

---

[9] For example, what is at issue in "Don't prepare anything" appears to be the alternative senses one might communicate to a prospective host or hostess, one of which is "Don't go to any special trouble or make anything fancy or elaborate," the other of which is "Don't make anything at all, I can't eat." It is not apparent how a misunderstanding of one of these for the other would be accommodated in the formats mentioned in the text.

[10] For those (like one of the anonymous referees) who find the transcript's rendering of working-class New York Polish and black dialect inaccessible, I offer here a proto-stenographic version (in normalized spelling and punctuation, and omitting overlaps, interruptions, etc.) of the transcript:

James:	Wait a minute, I've got to run ahead. Dadgummit, this is at least fifty thousand dollars.
Vic	He got his God d- you got your thing today?
James·	And I don't give a damn what door because I've got it here.
Vic:	Did you get- Let me ask you this.
James:	((laughs))
Vic.	Did you get your thing today?
James:	What?
Vic:	Your thing. (0.6)
James:	My thing?
Vic:	Yeah.
James:	I keep my thing with me all the time
Mike:	No, no, no man
Vic.	I'm not talking about that.
Mike:	He means- he means- he means <u>that</u> thing.
James.	((laughs))
Vic:	Did you get your thing today.
James:	Yeah, I got it, yeah I got it, I know what you mean; I'm just kidding ((laughs)), I got it alright. etc.

Excerpt 10 (US, 47)

James: WAIDAminnit, I gotta run ahead. Dad(gummit), this
is, at least hh eh- fifty thousn' dolluh ⌈( )
Vic: ⌊He got
iz god d-ehh ⌈you got your thing tuhday?
James: ⌊En' I don' give a damn what (door
cu:z) ⌈I got it heah.
Vic: ⌊Did you get-
Vic: Lemma ⌈ask yih dis.
James: ⌊AHHH hah hah ⌈hah- yeh- heh-heh¹
XX Vic: ⌊Didju getchor thing tuhday,
James. Wha:t.
Vic: Your thing.
(0.6)
Mike: ( )
James: Mah thing?
Vic. Yea:h=
James: =I keeps my thing with me aw:l ⌈t h e t i⌈m e.
Mike: ⌊No·, no no⌊ (man)
→   Vic: ⌊ I'm, =
= not ⌈talkin about dat
Mike: ⌊He means- he means-
James: ⌊AHH hah hah hah!
Mike: He means dat ⌈(ti:ng.) nhinhh!
→   Vic: ⌊Di.dju getchor thing, t'da:⌈y,
James. ⌊Yeh I got it,
Vic: ⌈(Well, Oh Wow::   )!
James: ⌊Yeh I got it, I know whatchu mean I dus'
kidd'n(hh) eh⁻heh heh!· hh I got it (owrigh⌈t)
Vic: ⌊(Okay.)
James: A(hh)heh a'ri I got ⌈my thi::ng,
Rich. ⌊(What's this thing),
(waitaminnit.) ehh hah hah!
James· Ehh heh heh ⌈heh heh
Vic: ⌊Ta:x. Yihknow,
James: Yeh.

Although in this episode Mike undertakes third position repair on behalf
of Vic who is the speaker of the trouble source (marked by *XX* in the left
margin), our discussion is directed to Vic's effort at repair. Note at the
first arrow a rejection component ("I'm not talkin about that"), rejecting
the understanding of "your thing" displayed in James's "joke," and
then, at the second arrow, a resaying of the trouble source turn. Vic's
effort at doing "clearer repeat" shows up in the transcript in the
stretched sounds, the stress on "you" (in Di:dju) and the "extra punctua-
tion" intonation drop, marked by a comma after "thing" in the tran-
script. (See also excerpt 13 below for use of the formulated punctuation
"period"—as a method for "saying it clearer.")

This practice of "doing clearer repeat" is, however, quite infrequent in the data that I have examined. By far the most common format for this last component, the repair proper, is the repair marker "I mean" followed by one or more of four operations on the trouble source that are designed to recast the recipient's understanding and, as we shall see, to provide an opportunity for another—different—response in next turn.

One type of repair operation takes the form "I mean" plus a contrast with the understanding of the trouble source displayed in $T2$, which may have been made explicit in a preceding rejection component, as in excerpt 4 above or in 11 below, which is taken from a group therapy session with teenagers.

Excerpt 11 (GTS, V, 12–13)

Roger:	. . . it's always this uhm image of who I am, 'n what I want people to think I am.
	(0.2)
Dan·	And somehow it's unrelated to what's going on at the moment?
Roger·	Yeah. But t(h)ell me is everybody like that or am I just out of ⌜it.
Ken·	⌞I- Not to change the subject but-
Roger.	Well <u>don't</u> change ⌜the subject. Answer me.
→ Ken	⌞No I mea- I'm on the subject.
→	I'm on the subject. But- I-I mean "not to
→	interrupt you but-" uh a lotta times I'm sitting in class, I'll start- uh I could be listening to the teacher and my mind'll be <u>four</u> million miles away.
Roger:	That's got nothina do wid' it.
Ken:	No. I mean I'm thinking about someday what I'm going to be, an stuff like ⌜that-
Roger:	⌞Heh wh(hh)en I grow
	up! heh ⌜hhh hheh hhh hh
Ken:	⌞No, no or-or I could picture myself as being a- being a pilot of a big jet plane or some such- gut rot.

Note at the arrowed turns that Ken produces first a repair-initiating component ("no"), then begins the repair proper ("I mea-") which he interrupts to respond to Roger's responsive next turn in a variant of the agreement/acceptance component (here addressing himself to a complaint, rather than to the response to a complaint, with "I'm on the subject"), and then returns to the repair proper with "I mean" plus an idiom directly contrastive with the one in the trouble source, a turn-initial marker of divergence from immediately preceding talk, but a different one.

A second type of operation is the reformulation of the trouble source, a resaying of the same thing in different words, again framed by "I

mean." In excerpt 12, a radio call-in show host is speaking with a caller who has recently developed a phobia about driving across bridges.

Excerpt 12 (BC, Beige, 14; Radio call-in show)

Caller:	. . but- hh lately? I have fears a' driving over a bridge.
	((pause))
Caller:	A:nd uh seems I uh- just can't uh (sit)- if I hevuh haftuh cross a bridge I jus', don't (goan' make-uh- do the) trip at all.
Host:	Whaddiyuh afraid of
Caller:	I dun'kno:w, see uh
→ Host:	Well I mean waitam'n. What kind of fear izzit. 'R you afraid yer gunnuh drive off the e:dge? 'R you afraid thet uh yer gonnuh get hit while yer on it? =
Host:	= ⌈What.
Caller:	⌊Off the edge 'r somethin.

The repair-initiation component here takes the form of "Well" plus repair proper (see n. 6 above), and the initial operation involves reformulating "Whaddiyuh afraid of" as "What kind of fear izzit."

So also in excerpt 6 above, in which two different senses of "Don't prepare anything" appeared to be involved: the trouble-source speaker employs as her repair operation a reformulation, "So I mean if you uh have a cuppa coffee or something, I mean that uh that'll be fine. But uh othuh th'n that don't uh don't bothuh with anything else."

A third operation that prosecutors of third position repair may employ in accomplishing the repair proper may be termed "specification." Whereas the first operation introduced above involves using a contrast to the trouble source, and the second involves using a different way of saying the same thing, specification involves introducing candidate specifics included within the earlier formulation of the trouble source. Thus, in excerpt 12 above, after doing a reformulation with "What kind of fear izzit," the speaker goes on to offer specific candidates: " 'R you afraid you're gonnah drive off the edge?" and so on. In 13, recorded in a Civil Defense headquarters in the course of a major hurricane, the Civil Defense director (Lehroff) asks his chief engineer about the weather.

Excerpt 13 (CDHQ, I, 46–47)

Lehroff:	What is the weathuh. Out in that area now.
Zebrach:	No winds, er it's squalling, rain, the winds are probably out of the north,- west, at uh estimated gusts of uh sixty to sixty five miles an hour.
( )·	(Whew!)
Zebrach.	Sustained winds of about thirty five to forty five miles per hour. And uh anticipated duration,

→	Lehroff:	How is the wah- weather period outside. Is it-
→		rain(ing)? uh windy? or what?
	Zebrach:	('s what I <u>said</u>). 's windy?
	( ):	( )
	Zebrach:	An' it's raining.
	Lehroff:	S' an' it's raining.
	Zebrach:	An' it's raining . . .

After initially doing "repeating it more clearly" Lehroff provides specifications of the sort of questions he "meant" to be asking (i.e., in vernacular rather than technical terms): "Is it raining or windy or what?"

A fourth type of operation used by repairers to recast the trouble-source turn may be termed "explanation." Consider in this regard excerpt 6 above, in which Belle includes, as part of her repair, an account of how she will be spending her day. In 6 this is apparently used as a preliminary to a reformulation of the trouble source, "Don't prepare anything."

In 14, however, the explanation is not preliminary to another operation; it is the repair proper. Here, a caller to a talk show has been on "hold" and suddenly finds himself on the air. When he breaks off his initial utterance (just as he was about to introduce himself, in violation of the "ground rules" for such programs) with the exclamation "Oh boy," the radio personality chortles at what he apparently understands as a sign of the zest with which the caller is taking up his opportunity.

Excerpt 14 (BC)

	Host:	And now, dear hearts, let's go to the next call.
		Shall we?
	Host:	Good evening, WNBC,
	Caller:	Good evening, this is uh, oh boy.
	Host.	ehh heh heh hyah ⌐hyah!
→	Caller:	⌊No I was listening to the
→		commercial, and I'm just kinda- confused fer a
→		min⌐ute.
	Host.	⌊Sorry about that, it's a little rattling.

At the arrowed turn, the caller offers an explanation of his prior turn that recasts it away from something that could properly be a laugh source, and, indeed, the host then offers a very different type of response, addressed to the type of turn which the trouble source has now been proposed to have been.

One other type of operation needs to be mentioned, one that does not occur with the "I mean" marker and that, unusual for the domain of repair, appears to be type specific for misunderstandings on the serious/nonserious dimension (Schegloff, 1987a; for other possible trouble-type-sensitive repair forms, see Schegloff [1991b]). The operation in these cases involves "characterization" of the trouble-source turn (typically by forms like "I was just kidding") and withdrawal from the sequence in which

it was implicated. Excerpt 15 comes from the early moments of conversation between erstwhile close friends who have not spoken for a long time.

Excerpt 15 (TG, 7–13)
```
    Ava:    I wan'dah know if yih got a-uh.m wutchimicallit.
            A: pah(hh)king place th's mornin'. 'hh
    Bee:    A pa:rking place,
    Ava:    Mm hm,
            (0.4)
    Bee:    Whe:re.
    Ava:    t! Oh: just anypl(h)la(h)ce? =
→   Ava:    = I wz jus' kidding yuh.
```

Bee's responses to Ava's inquiry show her to have taken it seriously and to be trying (through a series of next-turn repair sequences) to understand it so as to provide an appropriate answer. In undertaking to address this claimable misunderstanding, Ava recharacterizes the trouble-source turn as nonserious. Shortly thereafter the sequence is exited and the talk is turned topically in another direction.

These, then, are the components out of which third-position-repair initiated turns are built: a turn-initial particle that initiates repair, a response (agreement/acceptance) component to certain types of preceding turns, a rejection component in which the trouble-source speaker (and repairer) formulates the problematic understanding that has engendered repair, and the repair proper, accomplished by "clearer repetition," by characterizing the trouble-source turn as nonserious if it was taken seriously, and, most commonly, by one or more of four types of operations (more may remain to be described)—contrast, reformulation, specification, and explanation—commonly framed by the repair marker "I mean." Not all of these components need be present; any of them (including the repair proper; see excerpt 7 above) may be absent in any particular instance.[11] But whichever components are employed on any occasion, they are virtually always arrayed in their canonical order, the order in which they have been presented here. (The exceptions are taken up below.)

There is evidence of two sorts that this ordering is not incidental, but is the result of speakers' orientations to get the components into this ordering. The first sort of evidence comes from occurrences in which a speaker who is constructing a third position repair has begun to produce the repair proper, and then turns to incorporate an "earlier" component into the turn. Rather than completing the repair proper and then producing an agreement/acceptance or rejection component, repairers self-interrupt the repair component in progress, do the component that they

[11] Compare the second discussion of excerpt 9 below.

mean to insert, and then return to re-begin the repair proper that had been broken off.

In excerpt 11a (taken from excerpt 11, above), Ken intervenes after Roger has asked a question.

Excerpt 11a

Roger.   Yeah. But t(h)ell me is everybody like that or am
I just out of ⌈it.

Ken: ⌊I- Not to change the subject but-

Roger·   Well <u>don't</u> change ⌈the subject. <u>Answer</u> me.

→ Ken: ⌊No I mea- I'm on the subject.

→   I'm on the subject. But- I-I mean "not to

→   interrupt you but-" uh a lotta times I'm sitting
in class, I'll start- uh I could be listening to
the teacher and my mind'll be <u>four</u> million miles
away.

Note that in the arrowed turn, Ken begins with a repair initiator and then begins the repair proper with "I mea-." He cuts off the repair marker and inserts a variant form of an agreement/acceptance component, addressing himself directly to Roger's preceding complaint. Having inserted that *B* component in its proper place, he then resumes; that is, he re-begins the repair proper.

In excerpt 16, a high school student has called a radio call-in show to offer his views on a previously discussed topic, the rise of juvenile delinquency in the suburbs.

Excerpt 16 (BC, Beige, 4–6)

Caller: Well I::, do <u>not</u> find this sho:cking becuz I think
I c'n <u>understan'</u> why it's <u>beek</u>- it <u>has</u> been
increasing in the suburbs. Becuz <u>these kids</u> don't
uh·· they <u>don't</u> ro:b fer the purpose of <u>obtaining</u>
the object <u>usually</u>.

Host.   Mmhm::,

Caller·   <u>Many</u> of 'em <u>jes'</u> do it as a way of uh.:, taking
some parta <u>their</u> ti:me, let's say yihknow en <u>itchy</u>
finger so to speak.

    ·

    · ((1 1/2 pages omitted, re: why kids have so much
    ·   time, problems of transportation, etc.))
    ·

Host: So that the <u>kid</u> who robs a store::, en gets away
by <u>car</u> didn' do it buhcuz he couldn' get to the the
<u>drive</u> in ⌈yuh know.⌉

→ Caller: ⌊N o.   ⌋I-I-I <u>reelize</u> that <u>too</u>. But I-

→   mean, I don't mean these <u>people</u> who're committing

→   m:major crimes, b't I mean jus' the uh y'know

→   school vandalisms yuh know like broken windows . . .

In the third position turn here (at the arrows), the caller starts with a

repair initiation, then does an agreement component that responds to the radio personality's response to the caller's apparent complaint on behalf of teenagers, and then begins the repair proper with "But I mean." However, he then abandons it, and inserts a rejection component ("I don't mean . . ."), before repeating the repair marker to re-begin the repair. Other instances work similarly.

These fragments show that the canonical ordering of the components of third position repair turns is not only the product of a first-order "natural" ordering, but is reinforced by a second-order procedure that overrides the in-course production of such turns, so as to have their components arranged in "proper order."[12]

The basis for this order is quite clear. The initiation component serves, as other repair initiators do (Jefferson 1974; Schegloff 1979a), to put its recipient on alert that what follows may not be more of whatever unit has been transpiring, but that the progression of the talk may be being interrupted for repair. It is there to do the job of initiation, and that requires initial position.

The D component does the repair proper, in many cases re-doing the trouble-source turn in some variant version. In any case, as will become clear below, it is built to provide for another opportunity to respond to the trouble source, as it is newly understood. It therefore requires final position in the turn, after which the recipient can take next turn for a new "response," if there is to be one. The rejection component, then, necessarily must come between these two. With respect to the relative positioning of a rejection component and an agreement/acceptance component (where one is to be used), an agreeing response to a prior turn—especially in the delicate matter of complaints—best comes as early as possible, and surely before the very sequential basis for the prior turn is subverted by rejecting the understanding on which it has been based.

Given this "rational" basis for the canonical ordering of the components, and the apparent operation of both first- and second-order practices for ensuring it, how shall we understand deviations from it? If speakers can interrupt components-in-progress in order to insert ones that "belong" earlier, what is to be made of cases in which they do not seem to do so?

Such cases provide the second type of evidence for the account that has been offered. For when we examine instances of "misordering," as well as some instances in which an especially common component is omitted, we find the actual form of the turn to be specially adapted to its local sequential context. The practices for constructing third-position-

[12] On first-order and second-order practices and organizations, see Sacks and Schegloff, (1979, p. 16) and Schegloff (1987b, p. 75).

repair turns are, then, not applied mechanically, but are employed to implement particular interactional tacks.

Consider, for example, excerpt 17. Frieda and Reuben have come to have dinner with old friend Kathy and her husband Dave. Kathy and Dave are in the academic world, Frieda and Reuben are "in real estate." Shortly after arrival at Kathy and Dave's apartment, the following sequence transpires.

Excerpt 17 (KC-4, 10)

Kathy:	You got all dressed <u>up</u>? just to see <u>us</u>?
Reuben:	Are you kidding?
	((pause))
Frieda:	(I'm not dressed up ⌜ ripped)
Reuben:	⌞I'm in my underform.
Frieda:	I'm all ripped.
Kathy:	Oh yeah
Frieda:	Yeah
Kathy:	(I can see the hole)
Frieda:	⌜(all over)
Reuben:	⌞( )
Reuben:	Don't you recognize my uniform?
→ Kathy:	Yes. No, I meant Frieda was (wearing) a fancy dress.

Kathy's inital turn in this sequence appears to be doing a complaint (albeit a "mock" complaint), taking the formality of her guests' dress as a possible sign of a changed state of their relationship. Reuben and Frieda collide in their respective moves to offer a defense. Frieda wins the turn to go first and downgrades the state of her dress. Then Reuben offers as an excuse that his suit is "a uniform." So at the arrowed turn Kathy is speaking after an excuse has been offered as a response to her complaint.

Note then that her third position repair contains three of the components described above: a repair initiation ("No"), an agreement component with a response to a complaint ("yes"), and a repair proper ("I meant Frieda . . ."). But the A and B components are in reversed order.

Then note that the turn after which she is talking is a question of the yes/no type ("Don't you recognize . . ."), and that a "no" answer would be understood as rejecting the excuse that the question proffers. So, just as one common token of recipiency, "uh huh" (the so-called backchannel signal or continuer), cannot be employed after a yes/no question because in that environment it would be heard as a "yes" answer, so here the usual turn-initial particle for third position repairs cannot be used in turn-initial position because it would be heard as a "no" answer to the question and hence as a rejection of the excuse. The inversion of the repair-initiation and agreement/acceptance components here is, then, not a slip-up by the speaker, nor counterevidence to the account that has

been offered, but an adaptation of the constructional practices that I have already described to this particular sequential locus.

A similar account may be understood to pertain to excerpt 9.

Excerpt 9 (GTS)

```
        Dan:     . . . See Al tends, it seems, to pull in one or two
                 individuals on his side (there). This is part of
                 his power drive, see. He's gotta pull in, he can't
                 quite do it on his own. Yet.
         Al:     W'l-
       Roger:    Well so do I.
→       Dan:     Yeah. ⌈I'm not criticizing, I mean we'll just uh=
       Roger:          ⌊Oh you wanna talk about him.
        Dan:     =look, let's just talk.
       Roger:    Alright.
```

Dan's arrowed turn contains three of the four components that have been described, but it lacks an *A* component, one of those which occurs most commonly. Here, as in the preceding case, the repairer is speaking after a response to a perceived complaint, and a turn-initial *no* is vulnerable to being heard as a rejection of that response. Here Dan does not invert the components. He omits the repair initiator altogether, beginning his turn with an agreement to the response to the complaint and then proceeding to provide for a reanalysis by his interlocutors of the trouble-source turn. Like the inversion of components, the omission of some components can be understood, therefore, as adaptations of the sequential practices of third-position-repair construction in the service of the interactional stance that the speaker means to take up. Inversion and omission of components can be ways of doing things or ways of avoiding doing things.[13]

Placement of Third Position Repairs

The account that has been offered of the recurrent form of third position repair has been based on turns appearing in the sequential position characterized in excerpt 3 above. But that schematic characterization, although offered as referring to *positions in a sequence,* has been treated as if it were also mapped onto *turns in a series*—as if the several positions in a sequence necessarily occurred in consecutive turns.

As it happens, most third position repairs are in fact found in the consecutive (or "serially") third turn. But one payoff of developing an account of the format of third position repairs is that we are in a position

---

[13] For example, if the rejection component is not part of a third-position-repair turn, if the putative misunderstanding is not overtly rejected, the effect may be a sort of "be that as it may" function (as, e.g., in excerpt 5).

to recognize its characteristic appearance wherever it occurs. It can then be observed that, although the vast majority of third position repairs are in the third turn after the trouble-source turn, not all of them are, and so they are surely not there *necessarily*. We can then ask, Are third position repairs that occur in other than the serially third turn placed arbitrarily, or is some other orderly placement in effect? If their placement is orderly, we can try to develop an account which will both describe where such repairs are if they are not in the serially third turn, *and* which will ground our understanding of the occurrence of most instances, which are in third turn.

Some instances of third position repair that are not in third turn appear quite straightforward. In excerpt 18, the dispatcher at the fire department is talking to someone who has called in a report.

Excerpt 18 (FD, IV, 66)

Dispatch:	Now what was that house number you said =
	=⌈you were-
Caller·	=⌊No phone. No.
Dispatch:	Sir?
Caller:	No phone at all.
→ Dispatch:	No I mean the uh house number, ⌈Y-
Caller:	⌊Thirdy eight
	oh one?
Dispatch:	Thirdy eight oh <u>one</u>.

Here it seems quite clear that the third position repair, marked by the arrow, is displaced from serially third turn by the intervention of a next-turn initiated repair sequence ("Sir?" "No phone at all") addressed to the "next turn," the one displaying the misunderstanding.

Much the same is the case in excerpt 10 above, except there the intervention of next-turn-repair-initiated (NTRI) sequences was directed at the trouble-source turn itself, and there were two of them—first "What" and its response, and then "Mah thing" and its response. In both of these episodes, the next-turn initiated repair sequences entirely account for the displacement of third position repair from the serially third turn.

However, the third position repair in excerpt 17 above is not in the serially third position, and the intervening turns between it and the trouble-source turn are not taken up with repair. Is there an account that can deal with the "ordinary" cases in third turn, the ones displaced by NTRI sequences, and instances that seem to be anomalies, as does excerpt 17.

There *is* a formulation that appears to do both jobs. Although it sounds complicated, it is quite straightforward. Third position repairs are done in *the turn after a turn containing an utterance analyzably built to be "next" to some prior.*

The point is this. Although most turns respond to the immediately preceding talk, speakers can construct turns to address themselves to much earlier talk, even to talk occurring days or weeks previously. Especially by use of quotation formats ("Last week you were saying . . ."), speakers can locate any past talk—and even imagined talk—as what their ensuing talk is addressed to, is built as "next" to, although not all such efforts require as powerful a resource as quotation.

With this point noted, we can see that any turn can be built to display that it is addressed to some prior, and can then be understood as possibly revealing its speaker's understanding of the earlier talk to which it is addressed. Then, after such a turn, the speaker of the earlier talk which is being "responded to" can address whatever problematic understanding the "responding talk" may reveal to be informing it.

So, in excerpt 17, after Frieda works through her response to Kathy's complaint, Reuben undertakes to offer his response. His effort has been delayed not by NTRI sequences, but by Frieda's response. Still, his "Don't you recognize my uniform?" is understood by Kathy as addressed to her earlier turn about getting "all dressed up," and so when she undertakes her third position repair, she is doing so *in a turn after a turn containing an utterance analyzably built to be next to some prior,* and it is that prior that is the trouble-source turn and to which Kathy's repair refers.

A similar sequential logic is operating in excerpt 16, in which the third position repair is also placed at some distance from its trouble source.

Excerpt 16 (BC, Beige, 4–6)

Caller:	Well I::, do not find this sho:cking becuz I think I c'n understan' why it's beek- it has been increasing in the suburbs. Becuz these kids don't uh.: they don't ro:b fer the purpose of obtaining the object usually.
Host·	Mmhm::,
Caller:	Many of 'em jes' do it as a way of uh::, taking some parta their ti.me, let's say yihknow en itchy finger so to speak.

· ((1 1/2 pages omitted, re: why kids have so much
· time, problems of transportation, etc.))

Host:	So that the kid who robs a store·:, en gets away by car didn' do it buhcuz he couldn' get to the the drive in ⌈yuh know.⌉
→ Caller.	⌊N o. ⌋I-I-I reelize that too. But I-
→	mean, I don't mean these people who're committing
→	m:major crimes, b't I mean jus' the uh y'know
→	school vandalisms yuh know like broken windows . . .

Note first the reference to "rob" in the teenage caller's long initial turn. There is much ensuing talk on a variety of related matters in what follows and in the page and a half which has been omitted, but no pursuit or explicit mention of the matter of "robbing." Then, in the first turn after the deletion, the radio host builds into his talk a reference to "the kid who robs a store" and thereby builds into the analyzable construction of his talk that it is addressed to a particular prior. The third position repair, however remote it may otherwise appear to be from the trouble-source turn that it operates on, is in the turn after a turn analyzably built to be next after some prior—*that* prior.

This same account deals with third position repairs displaced by NTRI sequences as well. In both 10 and 18, and in other such cases, when a responsive turn *is* finally produced and registered—whether the NTRI repair was directed to what later turns out to be the trouble-source turn for the third position repair (its $T1$) or to the response to it (its $T2$), the third position repair is in "the turn after a turn . . ."

Finally, the same account explicates how and why most third position repairs are found in the serially third turn. They are there because most turns containing utterances analyzably built to be next after some prior are in the turn *after that prior*. Consequently, repairs initiated in the turn after them are in serially third turn.

With this, we have completed an initial account of both the *position* and the *composition* of third position repair. Before further explicating the job it does, and what the alternatives to it are, it will be useful to introduce a further position for repair, one not discussed in earlier literature on repair and in some ways quite different from other forms of repair. It supplies a complement to third position repair in the more general locus of repair after next turn and needs therefore to be introduced as a resource.

Because of the efficiency of the organization of repair, the deeper into the repair space one goes—the "later" the position—the fewer trouble sources have gone undetected and have "survived" to have repair initiated on them there. By the time of fourth position, very few trouble sources have eluded repair, and it is hard to find many "specimens" for study. The discussion of fourth position repair is, therefore, based on only a handful of cases and should be treated with special caution.

FOURTH POSITION REPAIR

Position and Composition of Fourth Position Repair

The occasioning of fourth position repair is represented schematically in excerpt 19, as third position repair was represented in 3.

Excerpt 19
A· $T1$ (e.g., $Q1$)
B: $T2$ (e.g., $A1$)
A: $T3$ (e.g., $Q2$)
B: ←

Speaker A does a turn, one which will turn out to be a trouble source; to track it, let us say it is an initial question. If B had some problem with $T1$, she or he could initiate repair in next turn. But if B does not encounter a problem in grasping $T1$, she produces a sequentially impli- cated next turn, $T2$; if $T1$ was an initial question, then let us suppose that $T2$ is an answer to that question.

If A found that $T2$ displayed some problematic understanding of $T1$, then she could next employ a third position repair to address that prob- lem. But if A finds no problem in $T2$ itself, or in the understanding that $T2$ displays of $T1$, then A can do a next turn, $T3$, predicated on the preceding sequence; let us say A does some sort of contingent, or follow-up, question—a second question. Now it happens that when B hears that second question at $T3$, it displays that the answer at $T2$ had been predicated on a problematic understanding, indeed an incorrect understanding, of $T1$. So following $T3$, in fourth position, B undertakes to address this problem, by dealing with $T1$ and the understanding which it was accorded.

This abstract schema was not abstractly derived. It describes particu- lar empirical sequences, which we now must examine. Consider first excerpt 20, which occurs in a research organization just after the begin- ning of the year. Loes is the receptionist and keeper of supplies; Marty is a visiting researcher.

Excerpt 20 (EAS, FN)
     Marty: Loes, do you have a calendar,
      Loes: Yeah ((reaches for her desk calendar))
     Marty: Do you have one that hangs on the wall?·
→   Loes·   Oh, you <u>want</u> one.
     Marty: Yeah

Marty's question at $T1$ can be understood as doing either of at least two actions. It can be the vehicle for a request for a calendar, or it can request to borrow a calendar (e.g., to check the day on which some future date falls). Loes apparently understands $T1$ to be doing the latter of these actions, and in $T2$ she offers as her answer a compliance marker and initiates the appropriate action. At $T3$ Marty does a follow-up request (built as a "follow-up" by its use, e.g., of "one" to refer to calendar, requiring reference to his earlier utterance for the understanding of this

one).[14] This follow-up request triggers in Loes a reanalysis of $T1$, a reanalysis whose product she proffers at $T4$.

As with third position repair, the schema represents these positions as occurring in consecutive turns, and the episode in 20 in fact runs off that way. But this is not necessarily the case for fourth position repair, any more than it was for third position repair. In excerpt 21, for example, Colonel Lehroff, the director of civil defense, is calling the home of the manager of the municipal truck yard.

Excerpt 21 (CDHQ, 15; Openings, 299)

	Phil:	Hello?
	Lehroff:	Phil!
	Phil:	Yeh.
	Lehroff:	Josh Lehroff.
	Phil:	Yeh.
	Lehroff:	Ah:: what've you gotten so far. Any requests to dispatch any trucks in any areas,
→	Phil:	Oh you want my daddy.
	Lehroff:	Yeah, Phi⌈l,
	Phil:	⌊Well he's outta town at a convention.

As it happens, the phone has been answered by the truck manager's son, whose name is also Phil. When he hears the caller's identification of him as "Phil," he understands himself to have been identified, and he ratifies the recognition (these being the $T1$ and $T2$ turns). The next two turns (the ones that end up displacing the fourth position repair from the serially fourth turn) complete the identification sequence by the caller's self-identification and the answerer's registering of it (Schegloff 1979b). These turns derive from a wholly independent order of organization—the organization of openings (as an aspect of the overall structural organization of single conversations, as here implemented in an identification se-

---

[14] Some have suggested that $T3$ here is itself a repair initiation, indeed a third-position-repair initiation, and that it has not been so treated because it fails to conform to the canonical format described earlier. The upshot of this suggestion is that the entire analysis is subverted because the canonical format is used to exclude nonconforming cases, which is how a canonical format is arrived at in the first instance. I have three responses. (1) Third position repairs that depart from the canonical format can be, and have been, recognized (see discussion of "I was just kidding" above; see also Schegloff [1991b]). (2) In $T3$, Marty is not so much addressing himself to trouble in Loes's understanding per se as he is using what her conduct displays her understanding to have been as the occasion for specifying his request. It is critical that he does not design his turn to display himself to be engaged in repair (e.g., "No, I mean, . . ."), but rather as a follow-up request, as noted in the text (3) Even were it the case that Marty was initiating third position repair on the understanding of which calendar he "meant," Loes's utterance is addressed to a *different* understanding problem—that he wants the calendar to keep (and not to borrow). This repair is addressed to the initial request, relative to which it is in fourth position, quite independent of the prior turn (*if,* again hypothetically, it *were* a repair initiation).

quence; see Schegloff 1986). The mutual identifications being completed, the caller proceeds to the first topic, a topic designed for the answerer whom the caller thinks he has identified and recognized and whom he had called to raise this very topic. But raising that topic (in *T*3, predicated on *T*1 and *T*2) reveals to Phil that his understanding of *T*1 was wrong, and his response at *T*2 was thus predicated on an incorrect understanding. In fourth position, which in this case is *not* in the serially fourth turn but is nonetheless clearly fourth *position,* he reveals the earlier misanalysis by announcing the product of a reanalysis.

In still other cases, fourth position is displaced from serially fourth turn by the interpolation of NTRI sequences, as was also the case for displaced third position repair (for the analysis of such a case, see Schegloff [1988*a*, pp. 59–60]).

As I have already suggested, the basic format of fourth position repair has two components. First is what Heritage (1984*b*) has termed the "change-of-state" token, "Oh." This is for the most part followed by a recharacterization of the *T*1, the trouble-source turn, as in "Oh, you want one," "Oh, you want my daddy," or (in the case referred to above and analyzed in Schegloff [1988*a*]) "Oh, you're looking for him." The proffered reanalysis is confirmed by the *T*1 speaker, and the repairer then offers a new response to the *T*1. In one case, the proffered reanalysis is omitted, and the fourth position repair consists only of the "oh" and a revised response to the trouble-source turn. This is excerpt 22 below (which is also dealt with elsewhere [Schegloff 1988*a*]).

Excerpt 22 (Kraus dinner)

	Mother:	'z everybody (0.2) ⌈wash for dinner?
	Gary:	⌊Yah
	Mother·	Daddy 'n I have t- both go ın different
		directions, en I wanna talk t'you about where I'm
		going (t'night).
	Russ:	Mm hm
	Gary:	Is it about us?
	Mother:	Uh huh
	Russ.	I know where yer goin,
	Mother:	W̅here.
	Russ:	To the uh (eighth grade) =
	Mother:	= Yeah. Right.
XX	Mother:	Do you know who's going to that meeting?
	Russ:	Who.
	Mother·	I don't kno:w.
→	Russ:	Oh·:. Prob'ly Missiz McOwen ('n detsa) en prob'ly
→		Missiz Cadry and some of the teachers.
		(0.4)
	Russ·	And the coun⌈sellors
	Mother:	⌊Missiz Cadry went to the- I'll tell
		you . . .

The trouble-source (marked by *XX* in the margin) lends itself to analysis either as a request for information or as a preannouncement. Russ initially understands it as a preannouncement and forwards the sequence to its next part. Mother's next turn reveals that this analysis is incorrect, and Russ then (in fourth position, which is the serially fourth turn) registers reanalysis without proffering its product (oh), and then produces a new response to *T*1 based on his reanalysis, in this case responding to it as a request for information by providing that information.

The Relationship of Third and Fourth Position Repair to Intersubjectivity

What third position repair is to the speaker of a trouble-source turn, fourth position repair is to its recipient's understanding of it. Third and fourth position are "self's" and "other's" (i.e., speaker's and recipient's) post-next-turn positions for dealing with problematic understandings of some turn (*T*1). This intimate, virtually mirror-image relationship of the two positions is evidenced in various ways.

Perhaps the most striking such evidence is the simultaneous occurrence of third and fourth position repair in fragment 9.

Excerpt 9 (GTS)
```
        Dan:     . . . See Al tends, it seems, to pull in one or two
                 individuals on his side (there). This is part of
                 his power drive, see. He's gotta pull in, he can't
                 quite do it on his own. Yet.
        Al:      W'l-
        Roger:   Well so do I.
→      Dan·     Yeah. ⌈I'm not criticizing, I mean we'll just uh =
→      Roger:          ⌊Oh you wanna talk about him.
        Dan:     = look, let's just talk
        Roger:   Alright.
```

We have already seen that in response to a perceived criticism of Al, Roger has asserted solidarity, and that at the first arrow Dan undertakes third position repair, omitting the repair initiation, thereby avoiding a seeming rejection of Roger's response to the criticism, but agreeing with it with an agreement/acceptance component, rejecting the understanding on which it was based in a rejection component, and offering a contrasting characterization in its stead.

Now note, at the second arrow, that Dan's agreement triggers a realization on Roger's part that his understanding of Dan's prior turn may have been faulty. At the same time as the speaker of that turn, Dan, is producing the critical and core remainder of his third position repair, Roger is producing a fourth position repair, in its standard format, "Oh,

you wanna talk about him." Third and fourth position are here used simultaneously and to do the same thing—third by speaker of the trouble source, fourth by its recipient.

We are now, finally, in a position to specify how it is that repair after next turn, composed of third and fourth position repair, provide the last structurally provided defense of intersubjectivity. These two positions are used to deal with *a sequence* going, or gone, off track. At least two turns are involved in the trouble—the trouble-source turn and the "next" turn, which displays a problematic understanding of the trouble-source turn. The major job that is served at third and fourth position is the retrieval of "next turn position" so as to allow another, better "fitted" next turn to be done.

Third position repairs accomplish this by the *speaker* redoing his or her own prior (trouble-source) turn, after which *its* next turn position is there again to be redone. And, regularly, it *is* redone and is done differently. A look at virtually any of the third-position-repair segments cited in the course of this discussion will show this (see, e.g., excerpts 4, 5, and 18).

Fourth position repairs allow retrieval of the next turn position by the recipient of the trouble-source turn by having the *recipient* reintroduce the trouble source *as reunderstood,* which, when confirmed, provides a new next turn position after it, in which a new—and different—next turn can be done, as in excerpts 18 ("Well he's outta town at a convention") or 22 ("prob'ly Missiz Cadry . . ." etc.).

Thus a sequence going awry because of a divergence between the parties' understandings of what is being done—a threatened breakdown of concerted action by virtue of an incipient loss of intersubjectivity—is caught at third or fourth position, and is set aright by realigning and meshing the parties' understandings and immediately embodying them in a succession of actions that, in effect, replaces an earlier succession and that resumes the interaction's course of action.

This is the sense in which these repair positions provide a defense of intersubjectivity. They are the last structurally provided positions because after these positions there is no systematic provision for catching divergent understandings. In general, after third position, such repair as gets initiated can at best be characterized as being initiated when the trouble source is "next relevant." Of course, it may never again be relevant.

Third and fourth positions provide that whenever some future turn addresses itself to some earlier talk and displays its speaker's understanding of that earlier talk, provision is thereby made for addressing trouble if there is any. And if a next turn predicated on a response to an earlier one reveals that that response was based on a problematic understanding,

provision is thereby made for retrieving that earlier turn and reanalyzing it, and re-responding to it based on the new analysis. These opportunities are engendered by the very fact of extending a sequence from some earlier talk. To the degree that further action is predicated on earlier talk, the understanding of that talk is in principle made available for review. The site of that review—composed of third and fourth positions—is the floating arena I have termed "repair after next turn," and is the last structurally in-built defense of intersubjectivity in conversation.[15]

THE MULTIPLE REPAIR SPACE

Schematic Display of Multiple Repair Space

It is possible, perhaps even natural, to read the preceding account as describing a set of objective places in conversation that determine or constrain the form of repair initiated in them: a speaker, finding himself or herself at some position, initiates repair of the type and format appropriate to that position.

But this would be an overly restrictive way of understanding the organization of talk in conversation, and perhaps of the relationship more generally between action on the one hand and structure or context on the other. In a way, one could as well say that the form of repair "selects" which "position" a speaker is adopting and choosing to speak from. A sense for the basis of this claim can be developed in the following manner.

Recall first that, if one takes any repairable as a point of reference, it is "followed" by a "repair initiation opportunity space" of four positions, generally lodged at least in the immediately consecutive turns (a turn's repair space can last "longer," of course, if some later talk is then ad-

[15] Some critical discussions of conversation analysis (e g , Taylor and Cameron 1987, pp. 120–23) have argued that analyses of the activities being done by a turn are condemned to an infinite regress, because "professional" analysis of any turn must be grounded in the understanding displayed by an interlocutor in next turn, but the import of that next turn must itself be grounded in its displayed understanding in the turn following, etc. Although there are a number of difficulties with this idea, what is most relevant here is that third and fourth repair positions provide a "floating" arena for catching and addressing any displayed misunderstandings, an arena renewed for any turn at talk whenever some subsequent turn addresses itself to that earlier talk. When an utterance has "cleared" this repair space without repair being initiated, it may be treated as having had the understandings displayed of it effectively ratified. While this is not a foolproof test (see the text below, where I discuss the contingency of a speaker *not* initiating repair in third position even when a clear misunderstanding has been displayed and see n. 19 below), the availability of displayed understandings to immediate repair initiation is a substantial and compelling basis for anchoring interpretations of utterances. It is so for professional analysts at least in part because it is so for conversational participants—at least until warrant is provided in the talk that occasions a reconsideration

dressed to it). That is, the following turns can have, as possible deployments of the talk in them, the initiation of next turn, third position or fourth position repair. These potentials must be understood to be there, to have been there, even if not activated (Schegloff, Jefferson, and Sacks 1977, pp. 374–75).

Recall next that nothing is excludable from the class "repairable," including repairs and repair initiators. That means that every turn trails a repair space behind it. And that means that there is not a single repair space, but that if there is ongoing talk, there are ongoing *multiple repair spaces*. Consider what this looks like represented in a schematic diagram.

Consider six consecutive turns, $T1$–$T6$. For convenience, I will treat $T1$ as the initial turn in its occasion; no talk has preceded it. I will treat the case of two interactants. For each turn, I will treat the production of some sequentially appropriate or sequentially implicated next turn as the central possibility, for simplicity alternating questions and answers, but will make explicit what alternatives are implied by the presence of multiple repair spaces. To minimize graphic complexity, I will not spell out all the possibilities.

$T1$	A:	$Q1$			
$T2$	B:	$A1$	NTRI ($T1$)		
$T3$	A:	$Q2$	NTRI ($T2$)	Repair 3d ($T1$)	
$T4$	B:	$A2$	NTRI ($T3$)	Repair 3d ($T2$)	Repair 4th ($T1$)
$T5$	A:	$Q3$	NTRI ($T4$)	Repair 3d ($T3$)	Repair 4th ($T2$)
$T6$	B:	$A3$	NTRI ($T5$)	Repair 3d ($T4$)	Repair 4th ($T3, 1$)

This diagram may be read as suggesting that, at $T2$, speaker B has an alternative to an answer as sequentially implicated next turn, namely an NTRI addressed to $T1$.

At $T3$, alternatives to a sequentially appropriate next turn (such as another question) include both an NTRI on $T2$ and a third position repair on $T1$.

At $T4$, alternatives to a sequentially implicated next turn, such as an answer (if $T3$ had been a question), include an NTRI on $T3$, a third position repair on $T2$, and a fourth position repair addressed to $T1$.

At $T5$, the same alternatives are available (with, of course, different trouble-source turns being addressed).

At $T6$, the same alternatives are still available, but now the fourth position repair option may be addressed not only to $T3$, but to $T1$.

Finally, if we remove the restriction that $T1$ be the first utterance of the conversation, $T1$, $T2$, and $T3$ would include the same complement of alternatives as the others, for they would have prior turns as potential trouble sources to which repair initiation could be addressed.

The result is that, as of the beginning of any turn position, it is, in

principle, a place for a sequentially appropriate next turn, *or* an NTRI, *or* a third position repair, *or* a fourth position repair.[16] Accordingly, "third position" or "fourth position," while being objective positions with determinate characteristics, are not fixed and are not positions in which speakers may "find themselves." One's position may be a sort of choice, turning in part on what is to be treated as the repairable—if, indeed, repair is to be initiated at all, which may also be an option. It will be helpful to explicate some of these claimed alternatives, choices and options empirically, and not merely as abstract possibilities in formal schemata.

Some Actual Alternatives

A first alternative to repair after next turn is this. A "next turn" that is treated as displaying a misunderstanding of prior can sometimes be seen in other ways—for example, as an interactional tack that its speaker is taking, such as a joke or wisecrack. In 10 above, for example, James's "I keeps mah thing with me all the time" is treated as based on a misunderstanding of prior turn, which warrants third position repair. It could as well be treated as a "joke first" response (Schegloff 1987a, pp. 212–16), to which the appropriate next turn is a laugh, after which a second, "serious" response may be expected. Thus, in excerpt 23, Ava has been describing her class schedule at college in response to an inquiry from Bee to the effect that "You're only in school late on Wednesdays." Then Bee, who attends adult continuing education, responds in light irony (for the sense of "early" and "late" is generically different for her school setting):

Excerpt 23 (TG, 434–42)
    Ava.   En then, the same thing is (uh) jus' tihday is
            like a long daycuz I have a break,
            (0.7)
    Bee: Hm:.
            (0.6)
    Bee: hh Not me:, hhuh uh-hhuh hhh! I go in late every
            day hh!

[16] Elsewhere (Schegloff et al. 1977, pp. 372–75), my colleagues and I showed that repair on each of several classes of repairable is initiated from each of the positions and also that repair on some same trouble-source token is initiated from the several positions. So, in principle, any utterance can have repair initiated on it from any of the positions. It does not follow, however, that every repairable can have repair initiated on it from every position, or that every turn position is a place from which each positional type of repair can be initiated. Particular trouble sources may not be addressable for particular types of trouble from each position, and particular turn "slots" may not, given what has immediately preceded them, lend themselves to positional repair types

→   Ava: Eyeh hh⌈h!
    Bee: ⌊No this'z- No I̲ have my early class tihday
              et four thi:rdy.

Rather than addressing the problematic understanding of what counts as
"early" and "late," Ava appreciates that a little joke is being made,
and offers a little laugh in response, after which Bee employs the "joke-
to-serious transitioning no" and offers a serious exchange account of her
schedule.

A first alternative to repair after next turn, then, is to see, instead of
a misunderstanding warranting repair initiation, a different tack being
taken, with its own sequentially appropriate next.

A second alternative is this. If "next turn" is understood as indeed
displaying a misunderstanding of its prior, then speaker of that prior—of
the trouble-source turn—need not initiate repair, but can "let it go," in
other words, he or she can treat the responsive turn as if it were sequen-
tially appropriate and correct. The misunderstood speaker may then later
redo the misunderstood talk of the trouble-source turn as a "new utter-
ance"; that is, do it not in the manner of "doing it again," but doing it
for "another first time," to use Harold Garfinkel's felicitous phrase.[17]
Here again, as in the first alternative described above, instead of repair,
the speaker of some problematically understood talk produces a sequen-
tially appropriate next turn.

Consider excerpt 24, in which two older women (sisters) are talking on
the phone ("Bud" being Ann's husband).

Excerpt 24 (NB)
        Ann: Well I tellyuh b- uh Bud might go back up t'the
                 boat, He's out ridin' a bike now en 'e thought
                 'eed ⌈go up'n getta pa:per.
        Betty·   ⌊Oh:::.
        Betty: Oh 'e wasn' going- 'e didn' go fishi-eh-deh
                 ⌈didn't go go:lfing then
        Ann: ⌊Oh I can't go-
        Ann: Huh. Oh God I can't go inna boat fer a lo:ng time.
                 'E siz "N̲o boating er no::,"
→   Betty: Awww.
        Ann: ⌈"-golf,"
→   Betty·   ⌊Bud wasn't playing golf?
        Ann: No.
→   Betty.   Oh::. . .

---

[17] In not stopping the action ongoing in the talk to deal with "trouble," this alternative
resembles what Jefferson (1987) terms "embedded correction," both of these practices
involving, then, ways of dealing with problematic talk without the apparatus of
"repair."

Note that Betty's first turn in this excerpt presents an inference touched off by Ann's prior—actually, ongoing—turn, an inference about Bud. But Ann's response reveals her to have heard that inference as about herself, and she responds with what appears to be a complaint that she has been told (whether by Bud or by a "he" who is her doctor) that she cannot go boating, and so on. In the following turn, then, Betty could deal with this displayed misunderstanding with a third position repair: "No, I don't mean you, I mean Bud." She does not. At the first arrowed turn, she does a sequentially appropriate response to a complaint—an expression of sympathy or condolence, "Awww." Then, at the second arrow, she "redoes" the trouble-source turn, but with modifications that allow it to come off as being said here for the first time, features fitted to its current sequential environment rather than the sequential environment of its prior saying. I will mention three of these.

Note first that the second saying drops a marker of sequential linkage to prior turn, a sort of "inference marker," with which the first saying ended: "then." Its retention in a saying of the turn that is no longer adjacent to the inference source could underscore the character of this saying as a repeat, and the source of the inference is no longer in prior turn.

Note second that the "topic" of the inquiry is referred to in the first saying as "he" and in the second saying as "Bud." Again, the first saying followed a turn in which Bud was referred to by name, and the pronoun is an appropriate way of marking subsequent reference to a same referent. In the second saying that sequential proximity has been lost, and in the intervening talk another referent has been mentioned (a "he" who may not be Bud but the doctor) to whom "he" might inappropriately be taken to refer.

Note finally that the first saying is initiated by "Oh," a particle which regularly marks a "change of state" (Heritage 1984b), which can serve as an interruption marker and initiator, and which here serves as a "touch-off" marker (as I believe Sacks termed it), displaying that what was just said in the turn-in-progress has touched off what the oh-speaker is about to say. The second saying is not initiated in the same manner, for, of course, it is not interrupting, and what it will be saying has not just been touched off. However, the change-of-state token is not omitted entirely. It is simply displaced from turn-initial position to third position in the question/answer sequence (at the third arrow), where it marks change-of-state.[18] The shift is fitted to the displacement of the inference from a position directly after its source to a later point in the talk.

---

[18] Note that "oh" marks an *interactional* event, not necessarily a *cognitive* one. By the time B says the utterance whose confirmation she acknowledges with "oh," she

The second alternative to repair after next turn, then, recognizes that there has indeed been a misunderstanding but adopts nonrepair ways of dealing with it.[19] One by-product of such a practice should be mentioned. A speaker who declines to initiate repair on such a misunderstood turn may in effect appear to have endorsed the understanding of it which has been displayed in the response. But on some occasions, participants will prefer that result to actually initiating repair.

If "next turn" cannot be understood as embodying some interactional tack which its speaker is taking but *is* understood as displaying a misunderstanding of the prior, trouble-source turn, and the speaker of the trouble-source turn *is* to initiate repair, there is still an issue of alternatives. The speaker of the misunderstood turn may be in a position to initiate either (1) third position repair on his or her own prior turn *or* (2) next-turn initiated repair on the prior turn, the one which stands in a problematic relationship to the misunderstood turn. Between these two alternatives, the preference for self-correction (were it to be operating here as it does between same-turn and next-turn repair) would issue in a preference for third position repair, for that form of repair involves the repairer initiating repair on his or her own talk rather than on the talk of another.

But there appears to be a constraint on doing third position repair, which if not met seems to make doing third position repair problematic. A speaker of some prior talk must understand some next turn after it— some recognizably "responsive" turn—well enough to appreciate, first, that it is based on a misunderstanding of that to which it meant to be responsive, and, second, what that misunderstanding is. Such a grasp of the "responsive" turn seems necessary for the prospective repairer to know how to design a relevant repair (including, e.g., what to reject in a rejection component).

It is striking that this constraint is regularly satisfied (see Schegloff 1968, p. 1082). Misunderstandings are not randomly wrong; there appear

---

has long since (in interactionally temporal terms) "known" what she here registers interactionally. Overt cases of such a divergence between the cognitive and the interactional should serve to alert investigators not to treat "oh" as necessarily marking *when* something has been learned or inferred.

[19] Of course, a misunderstood speaker may let the misunderstanding pass, and *not* redo the misunderstood utterance later either (see Heritage and Atkinson 1984, p. 14, n. 6). Thus, not all troubles, however clear and ripe for repair, are dealt with—either overtly or covertly. And the absence of repair is not a *guarantee* that a next turn reflected an acceptable understanding of prior. There may, of course, be other aspects of the talk—so-called hidden agendas—which elude the understanding of a recipient (by design or not), and these may not manifest themselves in any specific turn, nor may there be any interest on the speaker's part in repairing such failures to understand.

to be at least some systematic connections between what some talk is "designed" to do or to refer to, and what it will be understood as doing or referring to, *either* correctly/acceptably or not (cf. Schegloff [1968, p. 1082] on the systematic mistakability of certain forms of greeting for summoning and [1987a] for a range of other systematic misunderstanding types). And not only are the misunderstandings orderly, but they are apparently understandable to speakers of the misunderstood talk, who might have been taken to be so committed to the sense that their talk was designed to embody that they would be incapable of seeing what sensible alternatives were available in it.

Sometimes, however, misunderstood speakers do not grasp that they have been misunderstood, and they may then simply fail to understand talk designed to be responsive to their own. In that case, we do indeed find them employing NTRIs rather than third position repair, for the trouble source is then not their own prior talk, which they see has been misunderstood, but the other's talk, which they fail to understand.

In excerpt 25, B has called a radio call-in show to ask what remedy there might be for a suspended driving license.[20]

Excerpt 25 (BC, Beige, 20–21)
```
 1.  Host:   How long 'ave you had the suspensh'n.
 2.  Caller:  eh, since Febooary d't'ird.
 3.  Host:   Since February the third. You uh: wha'diyuh do,
 4.          fer a living.
 5.  Caller:  Eh::m I woik inna driving school.
 6.  Host:   Inna dri:ving school.
 7.  Caller:  Yeh. I spoke t'you many ti:mes.
 8.  Host:   Oh Yeah. You gottuh beautiful thing goin'.
 9.          Haven'tche
10.  Caller:  Yea⌈::h,
11.  Host:       ⌊You can'- You can' make a living.
12.  Caller:  No, I manage yih know, I go by bus, de fellas
13.          drive me over you know,
14.  Host.   Yeh but ha'di- whaddiyuh do et school.
15.          ((pause))
16.  Caller:  Excuse me?
17.  Host:   Whaddiyuh do et school.
```

[20] A few "ground rules" of this program are relevant to the understanding of this fragment. Callers are not allowed to identify themselves by name, and this is apparently a problem for those callers who call frequently and come to count themselves as acquaintances of a sort. They may find themselves starting to self-identify as a matter of "routine" (Schegloff 1986) before aborting the self-identification, as in excerpt 14 in the text above. They sometimes find ways of making themselves recognizable or find themselves recognized by some attribute recalled from an earlier call. This appears pertinent in excerpt 25 in the text, when A reacts sharply to the information that B works in a driving school, and B apparently takes it as having been "personally recognized"—"Yeh, I spoke to you many times."

```
18.  Caller.    Whaddiyuh mean " 'n school."
19.   Host:     Well you work ettuh driving school, ┌right?
20.  Caller:                                        └Yeh but I
21.             jus' go to the motor vehicle 'n awl that. I'm not
22.             en instructor ye┌t.
23.   Host·                     └Oh I see. Y'don' 'aftuh worry
24.             about that.
25.  Caller:    No,
26   Host.      Okay.
```

Here, mutual understanding seems to be disintegrating as early as lines 5–6. By line 11, the host seems to be displaying misunderstanding, but the caller does not seem to grasp this. He is therefore not in position to see that "What do you do at school?" is predicated on the puzzle of a driving instructor who has lost his license. Were the caller to grasp that this question reflects a misunderstanding of "I work in a driving school" (line 5), he could do a third position repair, for example, "No, I'm not an instructor, I just deliver messages." Without that grasp, he simply cannot understand what the host is asking him. When his NTRI (line 16, "Excuse me") elicits only a repetition of the trouble source, he can only initiate repair again on the same trouble source. Eventually, the nature of the trouble becomes available to him.

We can see here one of the senses in which various positional types of repair are alternative at some given turn slot. In excerpt 25, NTRI and third position repair are alternatives after "What do you do at school?" (line 14). But that does not necessarily mean that the recipient of that turn can choose which one to do. The constraints on doing a third position repair are not met here, and speaker B "can only" do an NTRI. Still, the character of the multiple repair space organizes and provides the terms of the alternatives that inform that sequential slot.

In excerpt 25, a recipient of a turn does not understand how it is a next turn, or what sort of next turn it is, to his own prior utterance. In third position repairs, the speaker of a turn does understand how an interlocutor's ensuing talk is designed to be a next, or responsive, turn but treats that as reflecting a misunderstanding of what is being responded to. In a fourth position repair, one who has responded to another's utterance comes to see from the consequence that that response engenders that the response was based on a wrong understanding of that which was being responded to. In each of these positions, one of the participants, at some point, "realizes" that the talk has gone off the track, that the participants are not operating with the "treatably same" understandings of what is being talked about and of what is being done through the talk. These are the "defenses of intersubjectivity," and they are structurally provided in the sense that the contingencies of "responding," and the turn-taking and sequence organizations of conversa-

tion, provide opportunities for assessing and addressing incipient divergences whenever a "response" is due or has been recognizably done *as a response*. It may be useful to examine one instance that exemplifies what can happen to an interaction in which none of these devices works to allow an incipient breakdown of intersubjectivity to be spotted and dealt with.

A Breakdown of Intersubjectivity

Excerpt 26 is taken from a radio call-in show, recorded in the late 1960s, while the Vietnam war was in progress. The discussion has turned to the topic of possible withdrawal. "A" is the radio "host" and "B" the caller. It will be necessary to provide rather a long excerpt.

Excerpt 26 (BC, Red, 103–6)

    A: Our settlement. Whatever it should be, will pertain to
         the action of our forces.
    B: So then we killed- uh we could still leave a vacuum
         there. Because if ⌐the South Vietnamese ⌐gov'mint-
    A: └We could,   └If we pulled
         out wrong? Yes.
    B: So if the Sie- uh South Vietnam gov'mint doesn't go
         along with ours. Then we just go uh say goodbye 'n we leave
         the shores. ⌐This's what- that⌐'ll accomplish ⌐(   ).
    A: └Ye-   └Uh,   └D'you have-
    A: D'you remember a thing called "the Korean War."
    B: Oh yes.
    A·   D'you remember the talks at Panmunjum,
    B: Uh (fuh) years. ⌐Yes.
    A.   └Alright, now I'm gunna tell you something
         thet I personally observed in Seoul during that period.
         There were two groups of people going up 'n down the street
         haranguing the crow.ds, doing their darndes'to sell the
         people on.:, no talks at Panmunjum. No peace with the eh:·
         with the North etcetretcetera. Checking with some friends of
         mine in Intellingence, I c'n tell you. Uh, which I did at
         the ti:me, I c'n tell you thet those two groups of people
         represented on the one hand, the North Koreans, an' on the
         other, Synman Rhee, an' the South Koreans. The Chinese and
         the Americans were locked knee deep in a confrontation in
         that wa·r, but neither the North or the South Koreans was
         willing yet—tuh let their allies quit fighting each
         other, in order to try an' get outta the thing.
    B: Those are certainly not the conditions that prevai:l
         ⌐now,
    A: └At one time, the Thirdy Seventh Div- the Thirdy Seventh
         Regiment, under Mike McKayliss, was drawn up, facing
         South Tuh keep Syngman Rhee from sending his people across
         the D.M.Z.

B·   Well yeh, Bu' that's not the conditions that prevail

a→   <u>now</u> there ⌜you know.

A: ⌞No bu' what I'm saying to you i:s, thet if
it <u>got</u> that m- <u>ba:d</u>, or it got <u>necessary</u>, <u>our</u> primary
concern, we have a concern <u>for</u> South Vietnam's

b→   territorial integrity which is why we're the:re. But
our primary concern regarding <u>our</u> personnel, <u>any</u>
military commander has that primary loyal⌜ty.

B: ⌞No? Are·n' we

c→   there because of U.N. uh — doctrine?

A: ⌜<u>No::</u>.

B·   ⌞<u>Aren't</u> we there under the ⌜the (   )-

d→ A: ⌞<u>Where</u> didju ever get <u>that</u>
cockeyed idea.

e→ B: Whaddya <u>mea:n</u>.

A.   U.N. doctrine.

B: <u>We're</u> there, representin' the U. <u>N</u>. No?

A: Wouldu- You go ask the U.N., you'll get laughed out.
No.

B: <u>We're</u> there because- of our interests

A: ⌜Yes.

B: ⌞We're not there wavin the U.N. flag?

A: We're- There's no U.N. flag <u>there</u>. Thet's not a United
Nations force. The United Nations has never taken a
single action on this.
((pause))

A: ⌜I-

B: ⌞No. I think (this ti::me)- I think you're <u>wrong</u>.

A: Sorry sir, I'd suggest yuh check yer facts.

B: I think y- I uh ⌜(   )

f→ A: ⌞I will refrain from telling you you
don' know what cher talking abou⌜t,

B: ⌞I ⌜wish you <u>would</u>

A: ⌞I just suggest you
⌜talk- you check yer facts.

B: ⌞I wish you <u>would</u>.

B: Because this's <u>what</u> I read in- in the <u>newspapers</u>.

        ⌜That we represent-

A: ⌞Well, then you been reading some <u>pretty ba:d</u>
newspapers.

B: ⌜We represent the U. <u>N</u>. there.

A: ⌞F'give me, but I gotta go.

A.   Sir, I would suggest thet if that's the case you switch
newspapers.

B: Well I hope I c'n call you ba:ck an' <u>correct</u> you.

A: L'k <u>you</u> check it out. 'n call me.

B.   I'll do ⌜so.

A: ⌞Okay?

B: I certainly <u>will</u>.

A: Mm <u>gu'night</u>.

After the radio host reads several commercial announcements and takes the next several calls, the caller represented in 26 is reconnected with the host as per excerpt 27 below.

Excerpt 27 (BC, Red, 156–58)
A. Good evening, WNBC
B· You asked me tuh call you back, an' I did
A· Yessir.
B: A::n' uh, My reference is the World Almanac, page seven
    hundred an' seven:tee⌈::n, nineteen sixty eight Almanac,
A: ⌊Mmmhmm?
A: Mmmhmm
B: -states thet we're there, under the U.N. d- command.
A: Where- Wai'min'. Waitamin'. Wai'min'. Wai'min. Now
    either we're talking about the U.- Oh yer talkin 'bout
    Vietnam,
B: I'm talkin about Korea.
A: Oh.
    ((pause))
A· Well now we go back an' do it all over again.
B· Well what- you asked me tuh check up on it, ⌈(   )
A: ⌊No.
    No, I thoughtchu were talkin about Vietnam. 'at's why I
    was arguin' with yuh.
B: Oh no
A: No, ⌈of course we were i-uh-
B: ⌊(   ) unilateral ⌈(   )
A: ⌊We are there now under
    U.N. auspices, although we committed our troops before
    we even bothered tuh check with the U.⌈N.,
B: ⌊Right. Right.
    That's all, then we misunderstood each ⌈other.
A: ⌊Yeah. My
    apologies.
B: Righto.
A: Right. Gu'night sir,

At arrows *a*, *b*, and *c* in 26, the interlocutors' discussion of both South Korea and South Vietnam makes use of a succession of "there" references in which they lose track of what each is talking about. That leads A to believe that B is asserting that the United States is in South Vietnam under U.N. auspices, whereas B is making that claim about the United States presence in South Korea. Neither catches that there has been an error in reference deployment or reference understanding,[21] each rather

[21] This is perhaps the clearest instance in this article of a misunderstanding based on what Sacks (in press) called, in an early (fall 1967) discussion of many of the matters discussed here, a trouble in the "tying techniques," i.e., the techniques by which one utterance is "tied to" (or relates itself to) another.

prepared to believe the other either ill-informed or ill-tempered. The misunderstanding leads to overt disagreement, to overt challenges and negative assessments (e.g., at arrow d), and eventually (at arrow f) to what amounts to an insult. In fact, before the call is closed, the parties are virtually shouting at one another. At the close of the call, there has been something tantamount to a rupture in the relationship, such as it was.

As it happens, the parties are once again put in contact with each other, and have another opportunity to sort out that there has been a misunderstanding. An apology is extended and accepted, the relationship (remote as it is) is restored, and the conversation is closed on an amicable note.

But this second chance (or nth chance) is an artifact of the mass media setting of the exchange.[22] Once the initial conversation had closed on a nasty note, the parties might well have never spoken again. This may appear to matter little when the participants are a caller to a call-in show and its host. But the interactional consequences of undermined intersubjectivity can be the same when the interlocutors are husband and wife, employer and employee, heads of state, and so on, with far more serious consequences for relationships, employment histories, and the like, including life itself (cf. Whalen, Zimmerman, and Whalen 1988).

CONCLUSION

When a source of misunderstanding escapes the multiple repair space, a whole institutional superstructure that is sustained through talk-in-interaction can be compromised. And since virtually anything in the talk can be such a source of misunderstanding, the potential for trouble for that institutional superstructure can be vast. It is against those systematic potentials for subversion of social order that repair after next turn is the last structurally provided defense.

[22] It is appropriate to remark that the account offered here has been for "conversation," and not necessarily for all speech-exchange systems or organized forms of talk in interaction. Because the organization of repair is mapped onto a turn-based organization of talk, variation in the setting or context, or anything that can involve some transformation of the turn-taking system by which the talk is organized and may well carry with it differences in the organization of repair, and, with them, changes in the provisions for defending intersubjectivity (see Button [1987] on the absence of third position repair from employment interviews). One instructive setting may well be formal performances and ceremonies. Novice performers are often instructed not to try to initiate repair on miscues in their performance, and many seasoned performers and civil and religious functionaries avoid addressing themselves to the fixing of "mistakes." This has, in part, to do with unavailability of sanctioned options (e.g., in next turns) for spectators to enter into the effort to repair trouble.

What then can be said about this little package to which the management of intersubjectivity appears to be entrusted? Several simple but general points can be registered.

The defense of intersubjectivity is *procedural,* putting tools in the hands of the participants for doing what needs to be done, with no "substantive" prespecification of what, for example, the trouble might be or what its sources or "causes," of what sort of work might need doing, and how it should be done.

Which is to say, as well, that it is *party administered.* One upshot of that feature is that adequacy of understanding and intersubjectivity is assessed not against some general criterion of meaning or efficacy (such as convergent paraphrase), and not by "external" analysts, but by the parties themselves, vis-à-vis the exigencies of the circumstances in which they find themselves. Ordinarily such adequacy is evidenced by the appropriateness (as assessed by the previously misunderstood party) of the revised response that the repair operation engenders. Surely this is the sort of thing that Schutz (1962, pp. 16, 27 ff.) and Garfinkel (1967, chap. 1) had in mind by the phrase "adequacy for . . . practical purposes."

The defense of intersubjectivity is *locally managed, locally adapted,* and *recipient designed.* That is to say, although always a present resource, it is invoked by parties as local circumstances make relevant and for problems as they arise for just those parties in just those circumstances (even, e.g., if those problems take forms that no one else but those parties could understand).

The defense of intersubjectivity is *interactional* and *sequential,* coordinating the parties' activities in achieving a joint understanding of what is going on and how those events might have been incipiently misunderstood. It is set into operation in a turn-by-turn metric at just the point at which problematic understanding appears incipiently consequential, as evidenced in the ostensibly interactionally responsive conduct of an interactional coparticipant. In this regard, both the interactionally "responsive" character of conversation and the presence of the "author" of what is responded to are critical.

All of this is to say that the locus of order here is not the individual (or some analytic version of the individual) nor any broadly formulated societal institution, but rather the *procedural infrastructure of interaction,* and, in particular, the practices of talking in conversation.

Although this is not the appropriate occasion for a detailed explication of the matter, it may be of interest to consider briefly the relationship of the position informing the present analysis (and *mode* of analysis) to strands of contemporary sociological theorizing that appear to be related to it.

In many of the above-mentioned respects the account offered here

appears to be fundamentally at odds with a stance such as the one adopted by Habermas. Although "communicative action" is at the heart of his theorizing (Habermas 1984, 1987), and has been from early on (1970), and although "understanding" is central to his view of it, his work is virtually devoid of any examination of actual communicative action in ordinary social life. He has, of course, insisted (1984) on the need for a preempirical pragmatics, if not an a priori one, presumably to serve as necessary leverage for a critical theory. But in appropriating contemporary speech-act theory (largely taken over from Searle 1969), he has contributed to the subversion of his own goals by relying on an analytic resource that in effect casts action as atomistic, individualistic, atemporal, asequential, and asocial.

Habermas is but one of the contemporary theorists who cast talk and interaction in a central role in the dramatic structure of their theories, while relying almost entirely on others' accounts of them. Habermas appropriates Austin and Searle; Collins and Giddens adopt Goffman. But it is becoming clear that, whatever their merits, none of these students of linguistic action and/or conduct in interaction is a reliable guide to the organization of action-in-interaction at the level of detail that increasingly seems relevant. Theories made to stand on such supports can quickly lose their grounding in reality.

In the juxtaposition of theories such as those of Habermas, Collins, or Giddens with the empirically grounded mode of inquiry presented here, one may legitimately reject the question of how such empirical results are to be mapped to the terms of such theories—a question often put to analyses like the one presented here. Instead—and to the degree that empirical work appears to have conveyed the organization of action and interaction relevant to the parties[23]—one should ask what grounds there are for continuing to take seriously theories whose analytic center of gravity is located elsewhere.[24] I have already pointed out (Schegloff 1987c, p. 229), "One can argue . . . that any discipline that takes the understanding of human action as its goal must be answerable to such microanalysis as seems to offer a rigorous account of the details of social action *in its own terms*. . . . Compatibility with the terms of a microanalysis adequate to the details of singular bits of interaction is a (perhaps *the*) major constraint on articulation with other orders of theorizing." It is such a microanalysis that I aim to advance in this article.

[23] By "relevant" I mean not "available to lay articulation" (not, then, what I take it Giddens [1984, p. 7] terms *discursive consciousness*), but rather "demonstrably oriented to in the actual conduct of the interaction."

[24] Unless, of course, they are the sort of theory for which evidence of this sort is not relevant in the first instance, in which case it surely makes no sense to juxtapose the two.

In all the respects mentioned earlier in reviewing the upshot of the preceding analysis, the device for the management and defense of intersubjectivity is of a piece with the organization of the activity in which it operates—ordinary conversation.

As it happens, this activity—conversation and its transformations into other forms of talk-in-interaction—is the vehicle through which a very great portion of the ordinary business of all the major social institutions (and the minor ones as well) gets addressed and accomplished. It is evident that much—even most—of the work of such institutions as the economy (in its several institutional contexts), the polity (in *its* several contexts), the institutions for population replacement (courtship, marriage, socialization, and education), the law, religion, social control, culture, and so forth, is accomplished in episodes of talk-in-interaction. A resource for the management of intersubjectivity for the activity of conversation and other forms of talk-in-interaction has, on that account alone therefore, a very broad provenance. But two sorts of observations may be offered about the bearing of this account with respect to intersubjectivity regarding nonconversational (or nontalk) occurrences and features of the social world.

On the one hand, much of the grasp of the world that informs the sentience and conduct of members of a society is, or can be, managed (crystallized, assessed, challenged, clarified, revised, confirmed, shared, reinforced, etc.) through talk, including objects of understanding that are not themselves talk. Any reference to an object, person, action, dream, fantasy, that is, anything real, or unreal but mentionable, and indeed anything understood to be presumed or presupposed by what is said or conveyed, can be made the object of talk—not necessarily a topic, but what the talk is understood to bear on. Thus in providing for the management of intersubjectivity in talk, provision is made as well for the management of intersubjectivity regarding whatever can enter into the talk.

On the other hand, from the account of the management of intersubjectivity in conversation, some guidelines are offered for the exploration of related practices in other domains.

First, if the practices are not indigenous to the practices of *talk,* then they should be anticipated to be indigenous to whatever other activities are involved, relative to which intersubjective understanding is under examination. Thus, for example, for work practices that require coordination but whose work environment interferes with ordinary communication through talk (e.g., high-rise construction riveting), mechanisms for intersubjectivity should be sought in whatever forms the practices of work take. If these are gestural (in the literal sense, not the adaptation of it as a metaphor in symbolic interactionism), for example, they may depart in various ways from the turn-taking practices that appear to

organize talk, and one might expect that a repair organization through which intersubjectivity is managed will be of a piece with that organization. Where those practices differ from those of conversation, one may ask whether vulnerability to losses of intersubjectivity, or to the failure to catch it, is enhanced, or whether defenses of intersubjectivity are provided for in other ways.

Second, the more general import transcends the specific concern with intersubjectivity discussed here. It is that activities and their organization can be, and should be, studied locally (that is, in the environments of their natural occurrence) and through the detailed examination of the indigenous practices through which it is (or they are) composed. And further, that much more of what composes the social world than has been imagined by most sociologists may be investigated in terms of activities and their procedural infrastructure. A whole domain of inquiry awaits.

APPENDIX A

A Note on the Organization of Repair

A bit of background on the notion "repair" may be of use here. Past work has given strong indications of a fundamental form of organization in talk-in-interaction that provides mechanisms for the participants to deal with an immense variety of troubles in speaking, hearing, or understanding the talk. These range from inability to access a word when needed or to articulate it properly, to transient problems in hearing (e.g., due to ambient noise), to variously based problems of understanding; the "variety of troubles" thus includes various classes of problems and a virtually unlimited array of "sources" or "causes." This "self-righting mechanism" that allows talk-in-interaction to keep itself going in the face of such "problems" we have termed the organization of *repair* (Schegloff et al. 1977).[25] A brief resumé of some main features of the organization of repair will provide the context for the specific concerns of this article.

In describing this organization of repair in talk-in-interaction, it has proved fruitful to discriminate between the *initiation* of efforts to deal with trouble (whatever the type of trouble), and the subsequent trajectory of such efforts to *success or failure*. Furthermore, there appear to be distinct differences between *repair initiated by the speaker* of the talk in which the trouble occurs, and *repair initiated by others*. Repair initiated

[25] As used in this context, the term "repair" is not addressed to all divergences or difficulties of understanding, only ones presented by the production and uptake of the talk itself.

by the speaker of trouble-source talk is initiated in various positions relative to that talk; virtually all repair undertaken by others than the speaker of the trouble source is initiated in the *turn after the one in which the trouble source occurred*. Further, whereas repair initiated by speakers of the trouble regularly proceeds directly to *solve the problem* if possible, repair initiated by others ordinarily restricts itself to *raising the problem*, that is, *initiating* the repair, but leaving it for the speaker of the trouble source to actually accomplish the repair (see Schegloff et al. [1977] for further elaboration). Third position repair is, then, to be understood by reference to this organizational context.

APPENDIX B

Transcription Conventions

A brief guide to a few of the conventions employed in the transcripts may help the reader in what appears to be a more forbidding undertaking than it actually is. It is apparent from the excerpts printed in this article that some effort is made to have the spelling of the words roughly indicate the manner of their production, and there is often, therefore, a departure from normal spelling. Otherwise:

→	Arrows in the margin point to the lines of transcript relevant to the point being made in the text.
( )	Empty parentheses indicate talk too obscure to transcribe. Letters inside such parentheses indicate the transcriber's best estimate of what is being said.
[	Left-side brackets indicate where overlapping talk begins.
]	Right-side brackets indicate where overlapping talk ends.
((points))	Words in double parentheses indicate comments about the talk, not transcriptions of it.
(0.8)	Numbers in parentheses indicate periods of silence, in tenths of a second.
:::	Colons indicate a lengthening of the sound just preceding them, proportional to the number of colons.
becau-	A hyphen indicates an abrupt cut off or self-interruption of the sound in progress indicated by the preceding letter(s) (the example here represents a self-interrupted "because").
<u>He</u> says	Underlining indicates stress or emphasis.

A fuller glossary of notational conventions can be found in other sources (see, esp., Sacks, Schegloff, and Jefferson 1974; Atkinson and Heritage 1984, pp. ix–xvii.)

REFERENCES

Atkinson, J. M., and J. C. Heritage, eds 1984. *Structures of Social Action*. Cambridge. Cambridge University Press.
Button, G. 1987*a*. "Moving Out of Closings " In *Talk and Social Organization*, edited by G. Button and J. R. E. Lee. Clevedon: Multilingual Matters.
———. 1987*b*. "Answers as Interactional Products: Two Sequential Practices Used in Interviews " *Social Psychology Quarterly* 50.160–71.
Cicourel, A. V. 1972 "Basic and Normative Rules in the Negotiation of Status and Role " In *Studies in Social Interaction*, edited by D. N Sudnow. New York: Free Press.
——— 1974. *Cognitive Sociology: Language and Meaning in Social Interaction*. New York: Free Press
Clark, H H., and E F. Schaefer. 1989. "Contributing to Discourse." *Cognitive Science* 13:259–94
Durkheim, E. 1915. *The Elementary Forms of the Religious Life* London: George Allen & Unwin.
Garfinkel, H. 1967. *Studies in Ethnomethodology*. Englewood Cliffs, N.J.. Prentice-Hall.
Giddens, A. 1984. *The Constitution of Society*. Berkeley and Los Angeles: University of California Press
Goffman, E 1983. "Felicity's Condition." *American Journal of Sociology* 89 (1): 1–53.
Goodwin, C. 1979 "The Interactive Construction of a Sentence in Natural Conversation " In *Everyday Language: Studies in Ethnomethodology*, edited by G. Psathas. New York: Irvington.
———. 1981. *Conversational Organization· Interaction Between Speakers and Hearers*. New York: Academic Press.
Gumperz, J. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press.
Habermas, J. 1970. "Toward a Theory of Communicative Competence." In *Recent Sociology, no. 2*, edited by H. P. Dreitzel New York: Macmillan.
———. 1984 *The Theory of Communicative Action*, Vol. 1: *Reason and the Rationalization of Society*, translated by Thomas McCarthy. Boston: Beacon Press.
———. 1987. *The Theory of Communicative Action*, Vol. 2 *Lifeworld and System: A Critique of Functionalist Reason*, translated by Thomas McCarthy Boston: Beacon Press.
Heritage, J. C 1984*a*. *Garfinkel and Ethnomethodology*. Cambridge: Polity.
———. 1984*b*. "A Change-of-State Token and Aspects of Its Sequential Placement." In *Structures of Social Action*, edited by J. M. Atkinson and J. C. Heritage. Cambridge: Cambridge University Press.
Heritage, J. C., and J. M. Atkinson. 1984. Introduction to *Structures of Social Action*, edited by J. M. Atkinson and J C. Heritage. Cambridge: Cambridge University Press.
Jefferson, G. 1974. "Error Correction as an Interactional Resource " *Language in Society* 2.181–99.
———. 1980. "On 'Trouble-Premonitory' Response to Inquiry " *Sociological Inquiry* 50:153–85.
———. 1987. "On Exposed and Embedded Correction in Conversation." In *Talk and*

Social Organization, edited by G. Button and J. R. E. Lee. Clevedon: Multilingual Matters.

Joas, H. 1985. *G. H. Mead: A Contemporary Re-examination of His Thought,* translated by Raymond Meyer. Cambridge: Polity.

Mead, G. H. 1938. *Philosophy of the Act,* edited by C. W Morris. Chicago. University of Chicago Press

———. 1962. *Mind, Self, and Society: From the Standpoint of a Social Behaviorist,* edited by C W. Morris. Chicago. University of Chicago Press.

Moerman, M. 1988. *Talking Culture: Ethnography and Conversation Analyis.* Philadelphia: University of Pennsylvania Press.

Moerman, M., and H. Sacks. 1988. "On 'Understanding' in the Analysis of Natural Conversation." In *Talking Culture· Ethnography and Conversation Analysis.* Philadelphia: University of Pennsylvania Press.

Parsons, T. 1937. *The Structure of Social Action.* Glencoe, Ill . Free Press.

Pollner, M. 1987. *Mundane Reason: Reality in Everyday and Sociological Discourse* Cambridge: Cambridge University Press.

Pomerantz, A. 1984. "Agreeing and Disagreeing with Assessments· Some Features of Preferred/Dispreferred Turn Shapes." In *Structures of Social Action,* edited by J. M. Atkinson and J C. Heritage. Cambridge: Cambridge University Press.

Sacks, H. 1975. "Everyone Has to Lie." In *Sociocultural Dimensions of Language Use,* edited by B. Blount and M. Sanches. New York: Academic Press.

———. 1987. "On the Preferences for Agreement and Contiguity in Sequences in Conversation " Pp. 54–69 in *Talk and Social Organization,* edited by G. Button and J. R. E. Lee. Clevedon: Multilingual Matters.

———. 1989. *Harvey Sacks—Lectures, 1964–1965,* edited by Gail Jefferson with an introduction by Emanuel A. Schegloff. *Human Studies* 12 (3, 4): 211–393.

———. In press. *Lectures on Conversation,* 2 vols. Edited by Gail Jefferson with introductions by Emanuel A. Schegloff. Oxford: Basil Blackwell.

Sacks, H., and E. A. Schegloff. 1979. "Two Preferences in the Organization of Reference to Persons in Conversation and Their Interaction " In *Everyday Language· Studies in Ethnomethodology,* edited by G. Psathas. New York: Irvington.

Sacks, H., E. A. Schegloff, and G Jefferson. 1974. "A Simplest Systematics for the Organization of Turntaking for Conversation." *Language* 50.696–735.

Schegloff, E. A. 1968. "Sequencing in Conversational Openings " *American Anthropologist* 70:1075–95.

———. 1979a. The Relevance of Repair to Syntax-for-Conversation." Pp 261–86 in *Syntax and Semantics 12· Discourse and Syntax,* edited by T. Givon. New York: Academic Press.

———. 1979b. "Identification and Recognition in Telephone Conversation Openings." Pp. 23–78 in *Everyday Language· Studies in Ethnomethodology,* edited by G. Psathas. New York: Irvington.

———. 1984. "On Some Questions and Ambiguities in Conversation." Pp. 28–52 in *Structures of Social Action,* edited by J. M. Atkinson and J. C. Heritage. Cambridge: Cambridge University Press.

———. 1986. "The Routine as Achievement." *Human Studies* 9:11–151

———. 1987a. "Some Sources of Misunderstanding in Talk-in-Interaction " *Linguistics* 25 (1). 201–18.

———. 1987b. "Recycled Turn Beginnings: A Precise Repair Mechanism in Conversation's Turn-Taking Organization." Pp 70–85 in *Talk and Social Organization,* edited by G. Button and J. R E. Lee. Clevedon: Multilingual Matters.

———. 1987c "Between Macro and Micro· Contexts and Other Connections." Pp. 207–34 in *The Micro-Macro Link,* edited by J. C. Alexander, B Giesen, R. Munch, and N J. Smelser. Berkeley and Los Angeles. University of California Press.

———. 1988a. "Presequences and Indirection." *Journal of Pragmatics* 12:55–62.

————— 1988*b*. "Goffman and the Analysis of Conversation." Pp. 89–135 in *Erving Goffman: Exploring the Interaction Order,* edited by P. Drew and A. Wootton. Cambridge: Polity.

—————. 1990. "On the Organization of Sequences as a Source of 'Coherence' in Talk-in-Interaction." Pp. 51–77 in *Conversational Coherence and Its Development,* edited by B. Dorval. Norwood, N J.: Ablex.

—————. 1991*a*. "Reflections on Talk and Social Structure." Pp. 44–70 in *Talk and Social Structure,* edited by D. Boden and D. Zimmerman. Cambridge: Polity.

—————. 1991*b*. "Conversation Analysis and Socially Shared Cognition." Pp. 150–71 in *Perspectives on Socially Shared Cognition,* edited by L. Resnick, J. Levine, and S. Behrend. Washington, D.C.. American Psychological Association.

—————. In press *a*. "On Talk and Its Institutional Occasions." In *Talk at Work,* edited by P Drew and J. C. Heritage. Cambridge: Cambridge University Press.

—————. In press *b*. "Third Turn Repair." In *Towards a Social Science of Language: A Festschrift for William Labov,* edited by G. R. Guy, J. Baugh, and D. Schiffrin. Cambridge: Cambridge University Press.

Schegloff, E. A., G Jefferson, and H. Sacks. 1977. "The Preference for Self Correction in the Organization of Repair in Conversation." *Language* 53:361–82.

Schegloff, E. A., and H. Sacks. 1973 "Opening Up Closings." *Semiotica* 7:289–327.

Schutz, A. [1932] 1967. *The Phenomenology of the Social World,* translated by G. Walsh and F. Lehnert. Evanston, Ill.: Northwestern University Press.

—————. 1962. *Collected Papers,* Vol. 1. Edited by Arvid Brodersen. The Hague: Martinus Nijhoff.

Searle, J. R. 1969. *Speech Acts.* Cambridge: Cambridge University Press.

Taylor, T. J., and D. Cameron. 1987. *Analyzing Conversation: Rules and Units in Discourse.* Oxford: Pergamon.

Whalen, J., D. H. Zimmerman, and M. R. Whalen. 1988. "When Words Fail: A Single Case Analysis." *Social Problems* 35 (4): 335–62.

Law That Does Not Fit Society: Sentencing Guidelines as a Neoclassical Reaction to the Dilemmas of Substantivized Law[1]

Joachim J. Savelsberg
University of Minnesota

Chances of realizing a legal rationality that does not fit society are limited. Referring to Weber's "Sociology of Law," to related themes in the sociology of polity, organizations, and occupations, and to recent debates on technocratization, juridification, delegalization, and responsive law, this article presents a theoretical discussion of this thesis. An empirical case, using the neoclassical concept of sentencing guidelines, exemplified by the federal and Minnesota cases, supports the argument. The neoclassical movement, aiming to reverse the substantivation of law and to correct lack of due process, functional failures, disparities, and discrimination, faces considerable impediments and may result in counterproductive effects. The article demonstrates that (1) societal conditions that caused substantivation hamper reformalization in the political process; organizational and occupational consequences of substantivation also impede (2) political and (3) implementation chances of neoclassical instruments; and (4) impediments to reformalization derive from the method of central guidance used to reestablish sociologically formal rationality.

I. INTRODUCTION: SOCIETY AND RATIONALITIES OF CRIMINAL LAW

Weber's writings on substantive rationalization deal with a sociolegal process characteristic of advanced stages of the modernization process. Substantive rationalization means the intrusion of economic, sociological,

and ethical criteria upon formal-rational reasoning and decision making (Weber 1978, esp. pp. 657 and 880–95). The term is closely related to contemporary debates on "juridification" (Teubner 1987, pp. 10–13), "technocratization" (Stryker 1989, 1990a, 1990b), "delegalization" (Abel 1982), and "responsive law" (Nonet and Selznick 1978).[2] Applying the term is both difficult and potentially fruitful in trying to understand (a) the development of law throughout the emergence and transformations of the interventionist welfare state and (b) the nature and consequences of current neoclassical countermovements.

Sentencing guidelines are a strategy of these countermovements in criminal law. They attempt to cure several dilemmas of substantivation that critics in sociolegal scholarship, the legal profession, the political sector, and the public have increasingly stressed: lack of due process, sentencing disparities, favoring the powerful. I argue and exemplify for the cases of the federal and Minnesota guidelines, for the policy-making and implementation processes, that neoclassical strategies meet serious structural and cultural impediments in modern society, conditions that caused substantive rationalization in the first place.[3]

Substantive rationalization is the most recent step in the historical process of legal rationalization and state development. The changing

[2] *Substantive rationalization* is related to, but distinct from each of these concepts. Weber's concept represents a crucial element of *juridification,* the process of increasing growth of law in 20th-century Western societies It is broader than *technocratization,* which means more specifically the application of *scientific expertise* to achieve substantive rationales My study of sentencing commissions shows that scientific expertise may also be used for attempts to reestablish formal rationality. *Delegalization* differs from substantivation when it refers to the replacement of legal regulations by nonlegal ones, but is one aspect of substantivation when it refers to the intrusion of extralegal rationales into legal decision making *Responsive* law is very similar to substantivized law. It differs only in that it contains a normative element that assumes the pursuit of substantive justice and provides safeguards against its misuse as repressive law.

[3] Interviewees on the federal case include five commissioners, five current and three former members of the research and directing staff of the commission, five members of the U.S Department of Justice, and several members of the staff of the U S Senate Committee on the Judiciary and the House Subcommittee on Criminal Justice Documents analyzed include transcripts of congressional hearings on the Sentencing Reform Act and on the sentencing guidelines, minutes of meetings of the Senate Committee on the Judiciary, papers presented before the Sentencing Commission, different drafts of the report and the supplementary report of this commission, and the Comprehensive Crime Control Act, including the Sentencing Reform Act.

rationality of law is based on the process of societal differentiation discussed in classical sociology (e.g., Durkheim, Parsons, Simmel, Weber) and, particularly in the late 19th and early 20th centuries, on changes in class relations and class conflicts (Therborn 1978; Turkel 1980). Positive and public bureaucratic law emerged with the differentiation of a specialized and centralized political sector. With the later separation of powers the law of the (now) republican state also became general and autonomous (Unger 1976, pp. 47–86). The result was described by Weber (1978) as formal-rational law (autonomous law in Nonet and Selznick [1978]). Because it created legal security and a stable foundation for economic exchange without intervening in societal structures, such law was basic to the bourgeois state and the capitalist economy. While it treated all persons as formally equal, it disregarded substantive social inequalities (e.g., Hurst 1956, p. 820). Under this law, to paraphrase Anatole France, it was forbidden to rich and poor alike to sleep under bridges and steal bread. In criminal law, the idea, prominently proposed by philosophers like Hegel and Kant (1965, pp. 99–106), was simply to be just, retributive, and to punish equal and free individuals equally for equal offenses.

It was recognized early that such formal-rational law contributes to social inequality (Marx and Engels 1938; Weber 1978, pp. 812–13). The resulting social, especially class, conflict was one of the strongest motors toward the interventionist welfare state and substantive-rational law (Alber 1982, pp. 200 ff.). Labor movements and political parties with labor constituencies demanded that those who create and apply the law ought to moderate the fiction of formal equality and liberty and acknowledge the substantively differentiated access to these desired goods (Ginsburgh 1979). In criminal law and related policy debates, concepts like intervention, therapy, rehabilitation, and social reform gained prominence. Other concerns also contributed to the eclipse of formal rationality. Weber (1978, p. 882) diagnosed "the desire to eliminate the formalities of normal legal procedure for the sake of a settlement that would be both expeditious and better adapted to the concrete case." He associates this "weakening of legal formalism" with a desire for "substantive expediency," a concern with impact, effect, and ends served by legal action, and adaptation "to the concrete case." This substantive rationalization of law results on the one hand in an intrusion of the state into society and, on the other hand, in an opening of state decision making to social ("extralegal") criteria (Teubner 1987, pp. 10–12).

Critics stressed follow-up problems of substantivation. Already Weber (1978, pp. 886–87 and 892–93) was concerned about an eventual loss of rationality and new problems to social justice. Particularly in most recent decades, and especially in the United States, substantivation was claimed

to be the source of counterproductive effects.[4] First, in the late 1960s and early 1970s, community and treatment programs began to evoke frustrations (these are most effectively expressed in Martinson's [1974] well-known "nothing works" resignation). In addition, a loss of due process was claimed to result from probation, indeterminate sentencing, parole, and preventive programs. Resignation and due process concerns were coupled with doubts concerning the justice of substantivized law. For example, the difficulties of clearly distinguishing between legal and extralegal determinants of sentencing (Nagel 1983) were related to the dilemma of high disparities in sentencing along the traditional set of legal criteria. In addition, substantive rationalization and related deformalization admitted both extralegal criteria and patterns of social inequality to legislative processes (Savelsberg and Brühl 1988; Savelsberg 1987, 1988) and legal decision making (Abel 1982). Numerous studies demonstrated that disparities in prosecution and sentencing tend to favor upper-class and powerful offenders (Benson and Walker 1988; Hagan and Parker 1885; Wheeler, Weisburd, and Bode 1982).

In the United States, frustration with substantivized criminal law legitimized a series of steps attempting to reverse the trend from formal to substantive rationality. The first step was a retreat to a general deterrence perspective (e.g., Wilson 1975), including a return to imprisonment and determinate sentences, that was expressed in many legislative efforts on the state level. As a result, prisons became overcrowded and many states could no longer meet costs. A decade of enforced imprisonment proved this strategy could not solve the crime problem (Pontell 1984). Rates of crime and recidivism remained high and continued to increase, resulting in further frustration and ironically in the incarceration of some who had caused jail overcrowding (Pontell et al. 1988). The response among some criminal lawyers was further withdrawal from general deterrence to a retributive, neoclassical just-desert principle (e.g., Von Hirsch 1987a, p. 85), combined with an attempt to control the prison population and to equalize sentencing practices through guidelines.

It is fascinating to observe how two decades of criminal justice policy have tried to reverse a long historical process leading from retribution to

---

[4] Other Western countries also experienced neoclassical movements. In the Federal Republic of Germany this movement is motivated and legitimized by constitutional guarantees of individual rights, similar to those of the U.S Constitution. This movement never grew as strong as its American counterpart. Two factors account for this difference. On the ideological side, the German constitution guarantees the welfare state in addition to individual rights (*Grundgesetz,* art. 20, 1) and thereby legitimizes interventionism. On the structural side, German neocorporatism, as opposed to American pluralist interest intermediation, secures welfare-oriented interventionism.

the substantivation of criminal law. The neoclassical movement does *not* seek ways to diminish the risks of substantive law while accepting its existence. It aims to redesign law, to clearly redraw the borders that had separated state and society and return to formal-rational law. Since the movement does not automatically get translated into legal practice, its members attempt to institutionalize reformalization through programs of central state guidance.

These neoclassical attempts offer an ideal case to confront the discussion on the autonomy or congruency of different subsystems or sectors in society, for example, the legal, political, social, and economic systems, with empirical reality. *I hypothesize that the chances to return to a past type of legal rationality are very limited.* More specifically, I argue that the sociostructural conditions that caused legal substantivation continue to be effective and that these conditions hinder the reconstruction of formal rationality. I argue further that substantivation itself has induced organizational and occupational changes that reinforce substantivation and now work as an additional force against reformalization in legislation and implementation. The interrelation between law, state formation, and forms of organization and occupation hypothesized here is not surprising if we understand law as that part of society's normative order backed by specialized agents of social control (Weber 1978, pp. 34–35), that is, in modern societies, formally organized state agencies that are staffed with specialized and professionalized personnel. In the theory discussion I further clarify Weber's concept of substantive rationalization, apply it to criminal law, and discuss its structural causes. A differentiated set of hypotheses is then followed by an empirical illustration for the case of sentencing guidelines.

II. SUBSTANTIVE RATIONALIZATION, ITS STRUCTURAL ROOTS AND CONSEQUENCES

One important clarification of Weber's "formal rationality" is warranted at the outset. Ewing (1987) demonstrates that Weber's theory distinguishes between (a) logically formal rationality in legal thought and (b) sociologically formal rationality of formal justice. Whereas the former refers to the generalization and systematization of legal subject matter so that it constitutes a deductive, logical, and gapless system of rules (see Schluchter 1981), the latter signals the rational-legal state emphasizing calculable enforcement of guaranteed rights. Sociologically formal rationality of formal justice "is abstract and bound by strict procedure, and guarantees the legal certainty essential for calculability in economic transactions, all of which applies to both civil [logically formal] and common

[not logically formal] law systems" (Ewing 1987, p. 489). Weber also includes the idea of noninterventionism when he points to the advantages of formal-rational law, as the basis for free societal and economic exchange, for the dominant classes (Weber 1978, pp. 812–13). Applying this conceptual clarification to our case, neoclassicism *aims at the reestablishment of sociologically formal rationality*. The *method* used for this end is central guidance, a logical and gapless system of rules, which provides for the *establishment of logically formal rationality*.

A. Substantive Rationalization of Criminal Law

Criminal law, in its quantity and quality (Black 1976) and in its relation to the political sector, varies with the societal context. Its long-term development, prominently discussed in sociological theory (Durkheim 1893; 1899/1900; Foucault 1975), contributes to criminal law's present dilemmas, which can best be understood as a result of the changing rationality of law. The process of substantive rationalization occurred in several ideal-typical stages, each of which represents a higher degree of substantive rationality (see table 1). First, legislators, legal scholars, and practitioners replaced vengeance by purposive action. "In criminal law, legal rationalization has replaced the purely mechanistic remedy of vengeance by rational 'ends of punishment' of an either ethical or utilitarian character, and has thereby increasingly introduced nonformal elements into legal practice" (Weber 1978, p. 884).

In the empirical development of criminal law, the principle of retribution was joined and partly overcome by that of general deterrence (see Bentham [1780] 1948, pp. 178–88). The idea was to deter offenders by increasing the anticipated cost of crime, that is, the risk of detection and punishment, beyond expected benefits. The purpose was no longer simply to do justice, but to reduce crime and to diminish its costs to society. General deterrence was still based on the punitive reaction to criminal acts, and the offender continued to be perceived as a freely acting, rational individual. But its strategy conflicted with the pure idea of just punishment, for example, by making public visibility of a case a criterion for punishment.

The next shift in the substantive rationalization of criminal law began in the 1870s, when rehabilitation rather than punishment became the primary purpose of criminal justice (Friedman 1985, pp. 595–600). In 1870 the National Prison Congress suggested indeterminate sentences as an appropriate tool toward this end (Samaha 1988, pp. 462–63). In the same year, the National Congress on Penitentiary and Reformatory Policy enthusiastically welcomed the new idea of parole (Samaha 1988, p.

TABLE 1

IDEAL-TYPICAL DEVELOPMENT MODEL OF CRIMINAL LAW AND JUSTICE

PURPOSES AND PRINCIPLES/MEANS/ STRATEGIES OF CRIMINAL LAW/ JUSTICE	TYPE OF LAW (Continuum)		
	Formal-Rational		Substantive-Rational
	Stage 1, Early 19th Century to Mid–19th Century	Stage 2, Mid–19th Century to Late 19th Century	Stage 3, Late 19th Century to Late 20th Century
Purpose(s)	Neutralize crime; reestablish equilibrium, simply to be just (Kant, Hegel)	Reduce the amount of crime; stabilize social order	Reduce the amount of crime; stabilize some aspects of social order, mobilize others, consider social justice
Principle(s)/mean(s) strategy(ies)	Retribution	General deterrence	Treatment of individuals (through psychology) and reform of communities (through social work, social reform, and preventive intervention)
Intervention in social order?	No	No, except through deterrent reaction to crime	Yes, through social work and treatment
Consideration of social justice in legal decisions?	No	No	Yes, by considering social/psychological restrictions to free will and pressures promoting offenses against legal order
Consideration of nonlegal consequences of legal decisions?	No	No	Yes, by considering ethical, social, and economic consequences for individual/family/ firm/community/ economy/society

688). Crime was no longer perceived as the act of freely and rationally acting individuals but as the act of persons restricted by social and psychological conditions. Social structure itself was no longer perceived as neutral but as causing the amount and kind of crime in society. This shift had three important consequences.

First, the objective of crime prevention now demanded a change of social and individual conditions through, for example, social work and psychological and psychiatric treatment. A clear expression of this process is indeterminate sentencing, which by 1922 had been introduced by the legislatures of 37 (of the then 48) American states. In addition, 44 states had created parole boards (Feeley 1983, p. 116). The end of a prison term no longer depended on the amount of guilt assigned to the offender, but on the improvement of the inmate as determined by social workers and psychologists rather than by lawyers. Finally, preventive intervention by police or social workers clearly indicated criminal law's shift away from responses to guilt and toward a very complex set of intermediary substantive purposes.

Second, to the degree to which guilt continued to be determined and considered, the substantive social and psychological conditions of offenders were increasingly considered. Individuals were no longer regarded as equal in their chances to be law-abiding or lawbreaking. The state and the law were no longer blind to social conditions. Formal justice lost relevance to substantive justice. This is most visible in the development of American criminal justice. The substantivation movement, starting in progressive municipalities and states around the turn of the century, reaching the national level during the New Deal era, and arriving at its most recent peak during the Kennedy and Johnson administrations, aimed to supplement formal freedom and justice with substantive freedom and justice. The Supreme Court under Chief Justice Warren set liberal rulings (Friedman 1985, p. 686), which strengthened the defendant's position in the criminal process. In addition, the poor offender, segregated in the ghetto and suffering from systematic social discrimination, was recognized as a victim of society who should not be fully blamed for his or her offenses. The substantivation of criminal law is obviously closely related to welfare-oriented movements and, most recently, to the successes of the civil rights movement.

Third, and most far-reaching, the consequences of criminal justice legislation and sentencing were increasingly considered. Such repercussions may be (1) economic, as when bankruptcies occur as a consequence of criminal legislation and sentencing (Savelsberg and Brühl 1988, pp. 191–288; Savelsberg 1987, pp. 546–54; Wheeler, Mann, and Sarat 1988, pp. 124–65), (2) social, such as deviant labeling or social isolation of individuals and their families, or (3) ethical. They may be completely

independent of justice criteria. They represent an intrusion of substantive rationality upon formal legal reasoning that went far beyond the substantive criterion of crime prevention previously incorporated in the criminal justice system.

B.   Structural Conditions of the Substantive Rationalization of Law

It is essential to my argument that the process of substantive rationalization came about because of changing societal constraints that pushed the actions and decisions of legislators and judges toward the substantive rationalization of law. Changing structural, cultural, and organizational conditions between the early 19th century and the late 20th century lead to dissatisfaction with formal-rational law, articulated by several carrier groups, and finally result in substantive rationalization in all Western societies. I discuss them and specify for the American case, along the dimensions of pluralism versus neocorporatism, and case law versus code law.

1.   *Social structure and culture.*—First, the organization of labor shifted the balance of power from the capitalist class to the working class. This shift occurred in most European countries in the late 19th century, but, in the United States, periods of massive immigration, from the mid- to late 1800s up to 1920, slowed this shifting process, so that it reached its peak only during the 1930s (Weir and Skocpol 1985, p. 137). Changing class structures were associated with dramatic changes in knowledge. The idea of humans as free and rational actors was replaced by that of humans as constrained by social and psychological conditions. The idea of society justly rewarding its members, who start off with equal chances, was replaced by that of a power-related distribution of resources and unequal chances. This changing knowledge structure was reflected in the institutionalization of the social sciences in universities and government administrations (for the U.S. case, see Stryker [1989, 1990a]; see Garland [1985] on criminology in late 19th-century Britain).

These changes in class relations and knowledge resulted in demands for a law that would recognize the contribution of formal rationality to social inequality (Weber 1978, pp. 812–15) and intervene in society to change the unequal distribution of power and resources. The new power of the working class, organized in unions and associated political parties, served as a basis to pursue these demands in political and legal decision making. Weber (1978, p. 886) sees "new demands for a 'social law' to be based upon such emotionally colored ethical postulates as 'justice' and 'human dignity,' . . . [and] directed against the pure business morality, having arisen with the modern class problem." The power base and organization of the working class continues to be, if weakened by the

decline of old industries, represented in established political parties. The idea of unequal chances in societies characterized by formal equality is now shared by other disadvantaged populations, such as racial minorities and women. In addition, the social scientific knowledge of substantive limitations to freedom is firmly institutionalized in the systems of education and science. Demands for a social law are not likely to be muted.

Second, changes in the economic sector also resulted in demands that opposed formal-rational law. A partial shift from free markets to the formation of trusts and the increasing dependence of political systems on large corporations and foreign capital results in demands for a law that emphasizes the reasonability of political and legal decision making more than justice. Examples can be found in economic law (Turkel [1980] for the United States) and in criminal law (Hagan and Parker [1985] for Canada).

Third, the functional differentiation of modern societies increased dramatically and resulted in large-scale configurations and very complex exchange networks. These in turn caused growth of external costs and use of the free-rider principle. Demands emerged for an interventionist law that would restrict external costs, provide compensation, and supply public goods (Olson 1965). The resulting public-sector production (Habermas 1975) and the need to guide productive forces (Habermas 1970, 1975) and to compensate the markets' dysfunctional side effects (Therborn 1978, Habermas 1975) created additional demands to integrate scientific-technical expertise into public administration. In the United States, scientific-technical expertise was built into the National Labor Relations Board (NLRB) in 1935, the Federal Trade Commission (FTC) in 1915, and the Social Security Board (SSB) in 1935 (Stryker 1989, 1990a, 1990b) at the cost of pure formal legal rationales. While the need for government involvement was realized in all Western societies, the acceptance of government intervention in concrete cases varied by country. Kelman (1981) documents and explains, in historical terms, the association of American self-assertive values with adversary institutions in the United States and of Swedish deferent values with accommodationist institutions of the Swedish welfare state. While adversarial attitudes and institutions hamper the straight implementation of substantive rationales by administrative authorities in the United States, they leave additional options for the negotiation of substantive criteria in adversary processes.

Fourth, functional differentiation was accompanied by the emergence of a diversity of interests and organized interest groups. As a consequence, political decision making became increasingly influenced by organized lobbying, while the voting mechanism lost relevance. This process of interest formation and intermediation, however, took different paths in European welfare states than the ones pursued in the United

States. The form it took probably influences the type, intensity, and stability of legal substantivation. Interest intermediation in the United States occurs through pluralism, in Europe through neocorporatism characterized by monopolies of interest representation, obligatory membership, partial state control over interest groups, and coresponsibility for public policy exercised by these interest groups (Schmitter 1982, p. 260). In addition, while American interest groups relate to the state through pressure politics, Europeans tend to do so through concertation (Schmitter 1982, p. 263). Policy-making in the United States is thus characterized by more conflictual processes, a more fragmented interest-group structure consisting of a much greater number of groups, and interest groups working from the outside in. There are less substantive and procedural norms developing over time that favor compromises. While lobby groups and their experts are less systematically integrated in policy-making in the United States than in Europe, lawyers and litigation are more highly valued.

While a strong involvement of interest groups and their experts always tends to push policy instruments toward substantive orientations, the model of pluralist pressure politics may, for several reasons, leave a relatively high chance to formal legal orientation. First, the activation of interest groups tends to be less reliable. Second, highly diverse interest groups are less powerful and more likely to neutralize each other. Substantivation is less institutionalized. The actual chances to push for specific substantive rationales depend on the particular policy area and on the case in question. The American type of interest mediation also leaves room for a relative variety of substantive interests to be reflected in different policy instruments and at different times. The flexibility of the American system is further strengthened by the nature of its legislative instruments and of its two-party system. Roth (1987, pp. 15 ff.) describes the American system as universalistic personalism and confronts it with the universalistic bureaucracies of Western European countries. Representatives in the American political system are more immediately dependent on their constituency as opposed to political parties and their platforms. The American system is therefore more open to policy changes.[5] Particular substantive rationales are less safely guarded in the American system. Which rationales prevail depends on which voter and lobby groups are able to articulate their interests and mobilize their resources in the face of a concrete case.

[5] This is true at least for such policy issues and strategies for which voter support can be expected. The decline of parties and party discipline in the United States has led to policy paralysis in other areas

Two final factors are of historical and contemporary relevance. First, lawyers feel their social status to be endangered by formal-rational law, which reduces them to "a slot machine into which one just drops the facts (plus the fee) in order to have it spew out the decision (plus opinion)" (Weber 1978, p. 886). This threat appeared particularly urgent in code law countries. Anglo-American case law always has made it relatively difficult to standardize decision-making criteria. In comparison with civil law countries, Anglo-American law, built from concrete situations in particular cases, leaves the judge with a relatively high degree of freedom to achieve a desired outcome. The Anglo-American lawyers are less of a driving force toward substantivation than their Continental colleagues, but they certainly act as a safeguard when centralized standardization of formal rationality is attempted. Second, law clients' discontent impedes logically and sociologically formal rationality. Clients want their cases to be dealt with not formalistically but reasonably, expeditiously, and with respect for the particularities of their own case (Weber 1978, p. 882). The general discontent of lawyers and clients is expressed by the diversity of organized interests that are introduced into the content of law through voting patterns and lobbying.

In sum, particularistic interest groups encourage substantive rationalization in the United States as well as in European societies. Their success is more reliable in the latter and further ensured by differences in parliamentary organization and the design of electoral systems. Yet, in both types of society these changes in social structure and interest-group formation lead to demands for an interventionist, purpose- and often welfare-oriented type of law that aims at reasonability and adjusting legal decisions to individual cases. Notwithstanding national specifics, social structural change led away from formal-rational law to substantive rationality.

2. *Organizational conditions of substantive rationalization.*—New sociostructural conditions and persistent group interests, associated dissatisfactions, and the resulting substantive rationalization of law are closely interwoven with and expressed in the structure and substance of law, the legal profession, political discourse on law, and the organizational contexts in which legal decisions are made and implemented. While impeding but—particularly in the American case—not fully excluding the political, legislative, or administrative construction of programs with neoclassical implications, these factors constitute an *implementation structure* that is hostile to neoclassicism. Nonet and Selznick (1978, p. 22), in a similar argument, associate responsive law with a postbureaucratic organization structure.

First, the desire to adapt state programs to specific types of cases makes it difficult to define general rules and procedures. In Continental law, for

example, in the Federal Republic of Germany, more and more criminal law provisions are found in trade, securities, or labor law than in the general criminal code. This program differentiation is a step toward substantivation and is repeated in the United States on the organizational level. Special court chambers, prosecutors, police units, and administrations are created for specific types of crime—for example, economic crime. Second, and closely related, the desire to adapt state reactions to individual needs is expressed in a growing special preventive and treatment orientation in criminal law. This first became evident in the juvenile court system in the late 19th century and later in the general system. Third, case and treatment orientation has resulted in the increasing importance of professionals in the making and executing of legal decisions. In criminal law, the role of psychologists, psychiatrists, and social workers in decisions and programs on punishment, treatment, and parole has been growing in importance (see Rosner and Harmon 1989). Fourth, the growing weight of professionals no longer allows the legislature precisely to predetermine proper means. The state thus moves away from conditional programs, which inform decision makers under what conditions they must use clearly defined means (e.g., welfare service or punishment), toward purposive programs, which define the purpose (e.g., cure or rehabilitation) and leave the choice of the proper means (e.g., type and length of treatment) to the now-professionalized staff (Luhmann 1985). Fifth, all the changes discussed above result in a general loss of bureaucratic, hierarchical order in formal organizations (Fine 1984; Strauss 1978). In addition, organizational boundary spanners, charged with coordinating the actions of their own organization with those of others in increasingly dense interorganizational networks (e.g., Benson 1975), need increasing leeway for negotiation and thus are freed from the immediate control of the hierarchy (Luhmann 1976, p. 220). Bargaining intrudes into previously very positivistic and bureaucratized legal systems as three examples from the Federal Republic of Germany indicate: (1) recent instruction programs on bargaining in criminal courts, offered to defense attorneys, were fully booked; (2) in 1975 the procedural code increased the discretion of prosecutors (*Strafprozessordnung,* n.d., para. 153a); and (3) diversion programs, which settle cases prior to the formal court process, spread rapidly.

Finally, the increasing diversity of cultural patterns, values, and norms in modern societies combines with positive law's permanent potential for change (Luhmann 1985, pp. 159 ff.) and thereby contributes to ongoing political struggles about criminalization and decriminalization of a diversity of behaviors—for example, abortion, drug use, and white-collar offenses. This constellation is likely to enhance further the perceived relativity and instrumental orientation of criminal law and the negotiability

of judicial decisions. It may contribute to the growing consideration of victims as actors in the criminal justice process (Edelherz and Geis 1974; Frehsee 1987) further increasing the negotiative or communicative order of criminal law.[6] The compensation of victims and their protection from further harm, sometimes inflicted by the pursuit of formal justice, became additional criteria of substantive reason in criminal justice.

Two counterarguments state that features of contemporary American society favor formal rationality. First, individual rights, based on the American Constitution and rooted in Kant's philosophical tradition, have become increasingly important in all areas of U.S. life. In criminal law the idea of victims' rights results in demands for a response that fits the severity of the violation, not the situation and personal circumstances of the violator (e.g., Murphy and Coleman 1990, p. 122; Gewirth 1978, pp. 292–300). This philosophical construction is itself one expression of the neoclassical movement. The practical relevance of "victims' rights," however, is strictly limited by the basic organizational and legal foundations of conventional criminal law. Opponents in criminal courts are not the plaintiff versus the defendant, but the defendant versus "the people." The interests of criminal justice agents are hardly identical with victim interests. Sometimes victims give painful testimony in court. In other criminal cases no victims can be identified (as in so-called victimless crime). Further, the result of criminal verdicts is not the payment of damages as in tort cases but a criminal penalty. Such punishment may even deprive victims of the potential payment of damages when it results in the defendant's loss of job and income. Philosophical constructions of victims' rights as the basis of criminal law are therefore not backed by the social practice of criminal law.[7] Second, it has been argued that the concept of individual responsibility is particularly strong in the United States, a situation exemplified by medical malpractice and tort and liability law (Landes and Posner 1987; Shavell 1987). Whenever individuals are thought responsible for their actions, formal rationality appears as an appropriate basis for legal responses. But the belief in individual

---

[6] Eder (1986), in the tradition of Habermas, analyzes the historical stages of growing communicative complexity of criminal justice: (1) the independence of the judge from his king, (2) the shift from torture to the argumentative identification of truth, (3) the admission of the defendant or her attorney as a communicating participant in the process, (4) the creation of the public defender, (5) the admission of the public to the process, (6) the formal division of the judge from the state (introducing the state's attorney).

[7] A radical development of victims' rights philosophies would lead to shifts from criminal law to tort law or other types of conflict settlement (Black 1989, pp 74–88). Reformalization under such a radical proposal would escape some of the structural and cultural restrictions discussed in this article Its political chances are obviously very limited at present.

responsibility has competitors also in America. The report of the President's Commission on Law Enforcement and the Administration of Justice (1967) most prominently expressed the competing belief in the social and structural sources of crime.

3. *Deductive conclusions: Hypotheses on neoclassicism.*—No matter what the sociological conditions for the emergence of the neoclassical movement are,[8] structural conditions drastically limit the chances of the reformalization movement to realize its political and legal models. Its supporters are unlikely to reach the intended results. The attempt to realize neoclassical ideas or concepts meets major challenges in the political and implementation process. *I hypothesize that neoclassical ideas get distorted in the political process in which the policy instrument is developed and become further distorted in the process of implementation.*

The distortion in the political process is likely for several reasons. First, some of the structural features of modern societies, distinct from conditions of the liberalistic early bourgeois state, oppose reformalization in the legislative process: the representation of working-class and minority-group interests, through lobbying organizations as well as through representatives and senators with respective constituencies; the institutionalization of behavioral, social scientific knowledge in government and academia; the need for public goods; the representation of corporate interests; and interests of legal clients and the legal profession. Second, organizational consequences of substantivation result in additional obstacles to neoclassical legislation. It is likely that judges and other members of the legal profession have at least partly internalized the substantive rationales they have applied for many decades. Psychiatrists, psychologists, and social workers have a vested interest in the maintenance of substantive rationales. These groups tend to resist reformalization. On the other hand, some ideological and institutional features of American society promote neoclassical ideologies and may help legitimate at least some neoclassical elements in policy innovations. Diverse substantive rationales may prevail due to pluralistic forms of interest intermediation and policy formation. Ideas of individual responsibilities and rights may also give more legitimacy to the development of neoclassical features of policy instruments than they would in other societies. Third,

[8] Cullen, Maakestad, and Cavender (1987, p. 18) see the source of the neoclassical movement in attempts by the legal profession to reestablish the legitimacy of criminal law "by reaffirming the formally rational character of law." A class-specific approach would stress interests of dominant classes in the disregard of social inequalities, formality in societal exchange. and a rejection of previous redistributive programs (neoconservatism). Greenberg and Humphries (1980) point to liberal scholars as initiators of neoclassicism whose ideas were then co-opted by representatives of conservative class interests.

these instruments, however, would have to face additional impediments in the implementation process. They would have to be implemented through networks of decision makers in complex administrative environments and by actors with internalized substantive rationales who often depend on constituencies with interests in substantive reason (see, e.g., Levin [1977] on judges in Pittsburgh). Fourth, implementors will oppose reformalization especially if attempted through central guidance, that is, if reformalization threatens to place restraints on their autonomy. Distortions of neoclassical instruments are thus likely to lead to unintended and counterproductive effects.

III. SENTENCING GUIDELINES. SOLVING PROBLEMS BY TURNING LAW BACK?

I illustrate and specify these theoretically developed hypotheses by confronting them with empirical cases.[9] I discuss the attempts of the federal government to create sentencing guidelines, and I look at guideline implementation in Minnesota, a very advanced state in criminal justice policy, where such instruments were created a decade earlier.

Sentencing guidelines are a neoclassical strategy. They aim at the solution of at least one of the follow-up problems of substantive rationality in criminal law: the problem of sentencing disparities. Their inventors want to achieve this goal by centrally guiding sentencing and by thus increasing its justice. Some *just* want to make it just. They demand a return to retribution and thus an exclusion of substantive rationales altogether. They plea for radically neoclassical guidelines. Others want to rationalize sentencing; that is, they wish to orient it toward a specific purpose or clearly defined hierarchy of purposes. This implies a reduction of substantive rationales and the withdrawal of discretion from individualized sentencing, thereby excluding from consideration the substantive complexity of individual cases. Sentencing guidelines are a fascinating experiment; by systematically trying to operationalize and practice a neoclassical legal concept, they confront it with empirical reality. My discussion of sentencing guidelines is based on the literature (von Hirsch, Knapp, and Tonry 1987; also Champion 1989; Nagel 1990) and on my empirical study of the U.S. Sentencing Commission (see n. 3 above). I briefly describe the structure, rationale, and neoclassical implications of sentencing guidelines. I then point to crucial problems that develop when guidelines met social reality in guideline creation and implementation.

---

[9] I avoid the term "test" because it is uncertain how representative the selected cases are. Future comparative case studies are needed. The term test is adequate, however, in the strictest sense of falsifying basic and overly generalized assumptions of neoclassical philosophies.

A. Sentencing Guidelines: Their Structure and Rationale

Sentencing guidelines can best be explained through the structurally important decisions that need to be made in the process of their construction. I summarize the decisions of the Minnesota Sentencing Commission, whose guidelines have the reputation of relative success. These decisions are reflected in figure 1, a sentencing grid that is constructed in three steps and from which the "correct" sentence can be read.

Step 1.—The main dimensions of the grid are determined. In the Minnesota case, the first dimension represents the seriousness of the present offense. It reflects the neoclassical principle of just deserts or retribution. This dimension, the core of formal-rational criminal law, was particularly fought for by such inventors of guidelines as von Hirsch (1987*a,* p. 89): "There is little that a commission can do to make certain that the crime rate will fall, since that rate is largely beyond the control of sentencing policy. . . . A rationale emphasizing desert has a different mission: that of scaling punishments so as to reflect differences in the reprehensibleness of the criminal conduct." This mission has no purposive orientation. Its proponents do not intend to intervene in the social order or to consider social justice or nonlegal consequences. It is thus opposed to substantive rationality.

The second dimension, the criminal history score, concerns the number of previous convictions, a factor understood to be a predictor for recidivism. Incapacitation is the underlying principle of this dimension. Incapacitation represents a substantive orientation that aims at the reduction of crime in society, even at the cost of justice, because not committed but predicted crime leads to imprisonment. In our case it represents the first departure from the original formalist understanding of guidelines.

Step 2.—The dispositional line (the bold stair-step line in fig. 1) determines for which cases a prison sentence is to be given, the so-called in-out decision. First the relative weight of the two dimensions in step 1 (the steepness of the line) must be determined. The Minnesota commission compromised between the idea of the inventors of guidelines, who favored the just-desert principle (ideally expressed in a horizontal dispositional line), and the criminal justice practitioners, who practiced the principle of incapacitation (expressed in a very steep dispositional line).[10] The Minnesota grid is therefore characterized by a formal, neoclassical

[10] Had the commission followed von Hirsch's argumentation, this line would be almost even, implying the rationality of the neoclassical just-desert principle. Only severe offenses would be punished with imprisonment. The sentencing practice in Minnesota prior to the introduction of guidelines resulted in a steep dispositional line It stressed "incapacitation" and gave little consideration to "just desert." The compromise concedes that successful implementation depends on the guidelines' acceptance by practitioners.

The dispositional line on Minnesota's grid

	Criminal History Score						
Seriousness of Conviction Offense	0	1	2	3	4	5	6 or more
10 (e.g., 2d-degree murder)							
9 (e.g., felony-murder)							
8 (e.g., rape)				IN			
7 (e.g., armed robbery)							
6 (e.g., burglary of occupied dwelling)							
5 (e.g., burglary of unoccupied dwelling)							
4 (e.g., nonresidential burglary)							
3 (e.g., theft of $250 to $2,500)			OUT				
2 (e.g., lesser forgeries)							
1 (e.g., marijuana possession)							

FIG. 1.—Scheme of sentencing grid of the Minnesota guidelines (from *The Sentencing Commission and Its Guidelines* by Andrew von Hirsch, Kay A. Knapp, and Michael Tonry, © 1987 by Andrew von Hirsch, Kay A. Knapp, and Michael Tonry; reprinted with permission of Northeastern University Press).

orientation, but modified in the political process of guideline construction by criminal justice actors and politicians who wanted to include the substantive rationale of crime reduction. Second, the degree of seriousness of an offense was determined for which a prison sentence is to be given (the level of the dispositional line) and, consequently, the amount of needed prison capacity.

Step 3.—Further decisions made by the commission concern the distribution of prison cells within the "in" decisions, that is, the differences between relatively mild and relatively severe punishment; the number and inclusivenesss of the values of the seriousness and criminal history scales; the association of all offenses to different grades on the seriousness scale; and the definitions of mitigating and aggravating factors as well as factors inadmissable for departures from the guidelines.

B. Problems of Guideline Construction: Political Reality and Constraints

The decisions needed to construct sentencing guidelines must be made in a complex and highly political process. The structures in which this process occurs determine how the idea of guidelines materializes. I intend to show that these structures are systemically related to the sociostructural, cultural, and organizational forces that are the basis for substantive rationalization. The chances that a formal-rational policy instrument would result from such structures are, therefore, at least limited. I document this for the definition of sentencing purposes as part of the construction of the U.S. sentencing guidelines (1984–87). I draw from three phases of this process: (1) the drafting of the Sentencing Reform Act, (2) the selection of commissioners, and (3) the drafting of guidelines by the commission.[11]

Drafting the act.—The legislative phase in the creation of federal guidelines resulted in the Sentencing Reform Act passed by Congress in 1984 under the aegis of the Comprehensive Crime Control Act. The Sentencing Reform Act determined that a commission must be created to develop sentencing guidelines. It also prescribed that the guidelines be numerical, binding for sentencing judges, and that sentences must not vary by more than 25% from the guideline value except for reasons to be determined by the commission. Finally, the act abolishes parole. One of its key titles relates directly to the distinction between formal and substantive rationality in criminal law. Title 28 of the U.S. code, chapter

---

[11] For other descriptions of this process, written by participants and not interest free, see Nagel (1990, pp. 899–932) and Robinson (1987).

58, paragraph 991 obliges the U.S. Sentencing Commission to

> assure the meeting of the purposes of sentencing as set forth in section 3553(a)(2) of title 18, United States Code. . . . [This section states that] . . . the court, in determining the particular sentence to be imposed, shall consider—(1) the nature and circumstances of the offense and the history and the characteristics of the defendant; (2) the need for the sentence imposed—(A) to reflect the seriousness of the offense, to promote respect for the law, and to provide just punishment for the offense; (B) to afford adequate deterrence to criminal conduct; (C) to protect the public from further crimes of the defendant; and (D) to provide the defendant with needed educational and vocational training, medical care, or other correctional treatment in the most effective manner.

Further on in chapter 58, items (3)–(7) demand the observation of the guidelines by judges, the avoidance of sentencing disparities, and restitution to victims of crime.

Some elements of the act are obviously devoted to a neoclassical agenda—for example, the principle of comparable and equal punishment and the abolition of parole to make sentences determinate. Yet, the definition of purposes includes all other purposes developed in the history of criminal law: special prevention, incapacitation, general deterrence, rehabilitation, and restitution. This has several consequences relevant for our theory problem. First, some of these purposes are contradictory to one another. These include: just punishment (punish according to guilt) versus general deterrence (punish according to visibility); general deterrence versus training and treatment (which may be attractive goods for potential offenders); and just punishment versus restitution (punishment may cut offenders off from resources that enable restitution). Choices between these conflicting goals are left to later stages of the guideline construction and implementation process. Second, the law demands that the history and characteristics of the defendants be considered when the sentence is determined. Along with the contradictory quality of the principles of punishment, this demand is opposed to the purpose of creating a standardized instrument that guides and equalizes sentencing and that is truthful and free from internal contradictions. Third, the consideration of the sociological conditions of crime ("history and characteristics of the defendant") and the principle of intervention ("training" and "treatment") are contradictory to neoclassicism since they represent substantive rationality (see table 1 above).

The result of the legislative process, as reflected in the principles of punishment, is due to systematic structural features of modern societies that stand in the way of a return to formal rationality. Such features include parliamentary representation of constituencies with interests in a substantive welfare orientation of law, for example, the working class

and deprived minorities. A most prominent example is the chair of the House Subcommittee on Criminal Justice, Congressman John Conyers (Detroit), who was central in the legislative process that resulted in the 1984 sentencing act. His skepticism concerning the application of universalistic, guilt-based sentencing provisions is implied in his comments on the investigatory basis of tax-related white-collar cases: "I am also wondering about when is a tax case a white collar crime. I have had so much trouble with the IRS, which has been sending people out to break down doors of doctors working in inner city neighborhoods who didn't pay their tax bill. . . . They are the ones who get the padlock on the door. They get their car snatched. . . . You don't do that with the big tax boys—you don't take a business building and take that away from anybody" (U.S. House of Representatives, Subcommittee on Criminal Justice 1988, p. 490).

Another structural factor potentially harmful to the neoclassical elements of the law is the variety of organized interest groups involved in the legislative process. Interest groups testified nearly 250 times at congressional hearings on the criminal code reform.[12] One obvious case of the substantivizing effect of lobbyists can be seen when Judge Tjoflat, speaking for the Judicial Conference of the United States, praised previous and recent efforts of the judicial branch to improve sentencing, in references to "sentencing institutes across the country" and "a sentencing information system" (U.S. Senate, Committee on the Judiciary 1983, p. 639). While this weakens the urgency of calls for guidelines, the Judicial Conference did not oppose the idea as such. Yet, an important modification was outlined in a bill proposed by the Judicial Conference, which recommended "a guideline system stuctured by a committee of the judicial conference rather than an independent commission (p. 640). . . . Four of them would be active judges; three would never have been judges . . . [whereas under the Senate bill] the President appoints seven, with no determination if any of them be a judge or not" (p. 641). While the judges did not prevent the creation of an independent commission, they achieved a partial success. The Sentencing Reform Act of 1984 contains the following provision: "At least three of the members [of the independent commission] shall be Federal judges selected after considering a list of six judges recommended to the President by the Judicial Conference

[12] Examples are the Associated Builders and Contractors, Inc , the American Civil Liberties Union, the Citizens' Commission on Human Rights, the International Brotherhood of Teamsters, and the National Rifle Association (see for these and others *Reform of the Federal Criminal Laws* [U.S Senate, Committee on the Judiciary 1977, 1979a, 1979b, 1981]) Evidence for the substantivizing effect of lobby groups in this and the following stages is suggestive Additional comparative studies are needed for proof.

of the United States" (title 28, chap. 58, para. 991). Apparently the judges succeeded in keeping some control over the following two steps of the political process. Given their interests in decision-making autonomy, this partial success weakens the neoclassical program.

Selecting the commissioners.—The idea that sentencing guidelines should not be created through legislative decision-making bodies but through a commission was first proposed by proponents of neoclassicism. They believed that a commission's work would benefit from freedom from political pressure and expertise (Von Hirsch 1987*b*, p. 6). This neoclassical idea, however, also appears to be idealistic and threatened primarily by the political process of appointing and confirming commissioners. The Sentencing Reform Act demands the establishment of a U.S. Sentencing Commission to be appointed by the president "after consultation with representatives of judges, prosecuting attorneys, defense attorneys, law enforcement officials, senior citizens, victims of crime, and others interested, . . . by and with the advice and consent of the Senate" (chap. 58, para. 991(a)).

The commission, constituted through this political process, included as voting members: three federal circuit court judges, one member of the U.S. Parole Commission, and three academics. Political rationale was reflected in the strong representation of appointed judges with political affiliations. Judge S. G. Breyer had been a long-time aide on the Senate Judiciary Committee to Senator Edward Kennedy (Democrat), Judge W. W. Wilkins had been similarly related to Senator Strom Thurmond (Republican), and Judge G. E. MacKinnon had ties to the chief justice of the U.S. Supreme Court. The academics, during the Senate Judiciary Confirmation hearings, were supported by members of Congress (see U.S. Senate, Committee on the Judiciary 1985, pp. 308, 310–311, 316–17). Their philosophical positions had been essential to their nomination. The responsible actors in the U.S. Department of Justice had consciously selected one academic commissioner known to be a proponent of a just-desert strategy, and two known to be followers of a general deterrence philosophy (according to my interview with Ronald L. Gainer, the U.S. assistant attorney general responsible for the selection [December 16, 1987]). The academics presented their positions at the Senate Judiciary Committee's confirmation hearings. Positions taken at this point represent a delicate blend of the candidates' opinions, their obligation to build on the Sentencing Reform Act, and the politically dominant tendencies in the Senate Judiciary Committee.[13] It is thus important to note how

[13] Confirmation hearings for government appointments are carefully managed by White House aides. Each appointee has a "handler" who works with him or her to map out responses to expected questions.

the nominees neglect or accentuate specific aspects of the broadly defined and compromised act. Asked by Senator Mathias about attitudes toward the purpose of sentencing, the economist Michael Block responded, "I have a rather simple notion of the function of punishment. I think that it is to minimize the social costs of crime and punishment" (U.S. Senate, Committee on the Judiciary 1985, p. 335). The sociologist Ilene Nagel answered the same question: "Rehabilitation as a purpose for sentencing did not serve us well. . . . Other than that . . . my philosophy is well reflected in the legislation, both in its combination of deterrence considerations and considerations for just and fair sentences, and with consideration for the reduction in unwarranted disparities and hopefully, an increase in certainty" (U.S. Senate, Committee on the Judiciary 1985, pp. 335–36).[14] The legal scholar Paul H. Robinson commented on the same issue: "Deciding for a particular case . . . my preference there is that the punishment of an individual case be fair and just. . . . If we impose a punishment that is a fair and just punishment for that person, that will have all those effects of deterring crime and incapacitating and everything else" (U.S. Senate, Committee on the Judiciary 1985, p. 336).

The neoclassical just-desert principle (Nagel and Robinson), the principles of general deterrence (Block and Nagel)–placed in the middle of our ideal-typical formal-substantive continuum–and incapacitation (Block) are particularly stressed by the three academic candidates for the guideline commission. The substantive criterion of rehabilitation is explicitly rejected. Despite the contradictions between candidates' approaches and the broad formulations of the Sentencing Reform Act, all three candidates were confirmed by the Senate Judiciary Committee.

In sum, the recruitment process reflects both the limits and relative chances of the neoclassical agenda. Limitations are demonstrated by the strong representation of judges in the commission, success by the selection of academic commissioners with predominantly neoclassical positions. This success seems to be due to four specific conditions. First, the nomination and confirmation of academic commissioners was much less exposed to public attention than other decisions. Second, decisions on the nomination and confirmation of academics were presented as technocratic rather than political. Third, the number of actors involved in this

---

[14] In her recent article on the federal sentencing guidelines, Nagel (1990, p 928) argues that "the Sentencing Reform Act was passed, at least in part, to make patently clear the rejection of the rehabilitative model   in favor of the new basis for sentencing— to punish, to promote respect for law, to deter, and to incapacitate." Here one of the most influential actors in this reform neglects the explicit demand of the act, which is "to provide the defendant with needed educational and vocational training, medical care, and other correctional treatment in the most effective manner" (sec. 3553(a)(2)(D)).

step was limited. Fourth, there was relatively strong ideological agreement among this limited set of actors with regard to neoclassical ideals, and there was little pressure from administrative or fiscal restraints to consider criteria of substantive efficacy at this juncture.

Drafting the guidelines.—Were the academics in the U.S. Sentencing Commission able to carry their moderate neoclassical intention into the actual instrument of sentencing guidelines? After all, they represented three of the four full-time commissioners. Or did the structural conditions around and within the commission prevent the construction of a neoclassical instrument?

Internal frictions between a "just-desert faction," led by Robinson, the legal scholar, and a "general deterrence faction," represented by Block and Nagel, the social scientists, resulted in temporary paralysis of the commission's work (on splits between lawyers and social scientists in commissions see Katzmann [1980, pp. 86–111]). Only after the rejection of two drafts, during the last months before the end of the term set by Congress, was a third draft developed and finally passed by the commission. Commissioner Judge Breyer, who had participated in the drafting of the Sentencing Reform Act in the Senate Judiciary Committee, was the primary drafter of these guidelines (this was confirmed in interviews with S. G. Breyer on March 4, 1988, and with several former staff members on December 16 and 17, 1987). This draft was pragmatic, with the deleterious effect that a clear definition or weighing of purposes was not attempted (U.S. Sentencing Commission 1987, p. 1.3). The drafters applied "the rule of minimal rationality" (Breyer, interview from March 4, 1988). The U.S. guidelines follow federal judges' past average sentencing practice, calculated—by the social scientist commissioners and staff members—through regression analyses of previous years' sentencing data. Two modifications of past practice are the relatively liberal handling of reasons for which judges may depart from the guidelines, and the relatively severe punishment of white-collar offenders. The other commissioners' votes were finally gained when the term set by Congress was about to run out and the commissioners were not willing to accept a failure of their mission (interview with Denis J. Hauptly, the former general counsel of the commission [December 17, 1987]). The final character of the guidelines differs from the academics' original neoclassical intentions. The political pragmatism of a federal judge, well acquainted with the complex mechanisms of political consensus formation through his drafting experience with the Senate Judiciary Committee, prevailed.

This pragmatism paid tribute to a diversity of lobby groups representing a multitude of interests at the hearings held by the guideline commission. The 209 persons giving testimony at 13 hearings between April 1986 and March 1987 represented mostly the professional groups concerned by

the sentencing reform: 33 judges, 26 prosecutors, 22 criminal defense lawyers and public defenders, 16 probation officers, 13 police officers, and 5 prison administrators. Lawyers from law firms (24) and from bar associations (12) were also represented. Forty-seven speakers gave testimony for a diversity of formal interest and lobby groups, such as the National Rifle Association (2), the NAACP (2), and victim advocates (6). It is uncertain to what degree this testimony actually influenced the commission's decision-making process. But the considerable amount of time invested in these hearings by the commission within the short time available for the drafting of the guidelines indicates the diverse legitimatory pressures it perceived.

Among these diverse interests, those articulated by judges are of particular relevance for our consideration. Most immediately concerned among the implementing agencies, judges represented the largest group during the hearings, and three judges were members of the commission, including the main drafter of the guidelines. We have seen that judges, concerned with decision autonomy and related status ideologies, are one of the main carrier groups of substantive rationalization (Weber 1978, p. 886). And, indeed, during the hearing before the House Subcommittee on Criminal Justice (U.S. House of Representatives, Subcommittee on Criminal Justice 1988), major factions from among the judges clearly indicated their dislike of binding sentencing guidelines. Judge Wiseman, claiming to give testimony for the 500-member Federal Judges Organization, argued: "I personally believe, and I think I speak for my judges in this respect, [that] these guidelines are highly mechanistic. They almost make an automaton out of the judge. You might as well replace me with a computer. . . . It eliminates humanity from the sentencing process" (U.S. House of Representatives, Subcommittee on Criminal Justice 1987, p. 201, and pp. 480 and 233 for comments by Judges Merritt and Sweet, respectively). Even the relatively open guidelines drafted under Judge Breyer, in which he incorporates previous practice and leaves room for departures, appeared unacceptable to most judges. Commissioner Nagel (1990, p. 938) reports "While the initial formulation of the departure standards was two-tiered and fairly stringent, extreme political pressure—both by those adverse to the more punitive guideline sentences—and by those judges who . . . wanted to return to greater flexibility—was exerted to weaken . . . the standard so as to make departures more accessible."

Corporate business, represented by the Business Roundtable, was another interest group that successfully intervened in the promulgation process. In 1989, the guideline commission, in cooperation with the Criminal Division of the U.S. Department of Justice, had developed corporate guidelines. These guidelines, clearly oriented toward a just-desert, retrib-

utive philosophy, provoked hefty protest from the Business Roundtable, which pointed out the problematic economic consequences of the law. Their protests caused the White House Domestic Policy Council and, pressed by the White House, the attorney general to intervene. As a consequence, the corporate guidelines were withdrawn. Their fate is to be decided in future revisions (information from interviews; see also the *Washington Post* 1990*a*, 1990*b*).

In sum, the transformation of the idea of sentencing guidelines into a policy instrument occurred under specific constraints that systematically deviated from original neoclassical intentions toward a substantive direction. Still, in both the federal and Minnesota cases a guideline instrument was developed. The assumption persists in both cases that guidelines can be strictly implemented. How *does* a "truncated" (Robinson 1987) neoclassical instrument function in the implementation process?

C. Implementation Problems: Further Pressures toward Substantivation

Structural constraints, as in the political phase, distort the remaining neoclassical elements of guidelines in the implementation process. Since time-series studies on the implementation of the federal guidelines are not yet available, I return to the Minnesota case. The Minnesota guidelines still incorporate neoclassical elements although the originally stronger neoclassical intentions had been watered down. The Minnesota case is considered a relatively successful example and an evaluation study over a series of years exists (Knapp 1987).[15] Several problems emerge.

First, following an initial success, problems are registered with regard to the purpose of unification of sentencing and its adaptation to the guidelines. In 1980, the last year before the guidelines went into effect, 17%–18% of the sentences departed from the values defined by the commission. In 1981, the first year of the guidelines, the departures were only 6.2%, certainly a success in terms of the commission's purposes. Over the following years, however, departures increased to 8.9% in 1983 and 9.9% in 1984, the last year included in the evaluation study. One reason for the newly increasing departures is of particular interest in this context. During the period under investigation the proportion of indicted sexual offenses increased dramatically. Judges believed that in these cases rehabilitative concerns should direct the quality of the sentence, an attitude that resulted in considerably less imprisonment than the guidelines de-

[15] This is the only time-series evaluation analysis of sentencing guidelines for all offenses. More analyses over longer periods of time would be desirable but are not available at this point.

manded. This position was shared by the commission although it refused, for political and symbolic reasons, to alter the guidelines to reflect the new practice. In a situation of growing concern about sexual offenses and resulting political sensitivity, the commissioners were unwilling to downgrade the guideline value. Instead, they compromised the purpose of the "truthfulness" of the sentencing system. Aiming at effective general deterrence, the commissioners accepted the discrepancy between norm and reality. This case shows that guideline implementation, like guideline construction, cannot function independently of a complex political context.

Second, the guideline commission in Minnesota had attempted to change the relation of punishment between different types of offenses. They judged that property offenses, with 47% imprisonment, had been relatively severely punished in relation to violent offenses, with 39% imprisonment (values for 1980). When rating the severity of different offenses in the sentencing grid, they ranked property offenses relatively low and violent offenses relatively high in relation to past sentencing practice. In this case too we observe success in the first year of implementation which was later defeated. The chances that property offenders would be convicted to prison in relation to violent offenders shifted to 37%/57% in the initiation phase (1981), but moved back toward its original relation of 43%/50% in 1982 and 50%/43% in 1983 and 1984. A look at the sentencing grid (see fig. 1 above) helps us to understand this defeat. Prosecutors disagreed with the reevaluation of the commission. They tried to reestablish the old relations between punishment for violent and property offenders. They succeeded by defining charges for property offenses in a more detailed way so that the criminal history score of property offenders added up faster. For property offenders, the principle of incapacitation went into effect just as it was before the guidelines.

This finding points to the crucial problem that sentencing guidelines disregard a basic sociological fact of modern organizations. Sentencing, like all decision making, is not an isolated event but an element in a chain of interconnected events. The intervention in sentencing decisions as one element of the criminal justice system results in possibly neutralizing reactions in other parts of that system. In more concrete terms, future research should test the hypothesis that sentencing guidelines increase the power of the prosecutor in the negotiations preceding the indictment. Since the sentence for particular offenses is clearly defined through the guidelines, negotiations about the sentences are no longer possible. The prosecutor can also no longer threaten with unrealistically high sentences. Yet, under certain circumstances, guidelines may allow her to increase pressure on the defendant to enter into charge bargaining. A shift from sentence to charge bargaining must be expected. The sentence considered

adequate for a certain offender with a given offense, for reasons documented in past sentencing research, will increasingly depend on charge and fact bargaining and be rationalized and legitimized through the definition of charges.[16]

Finally, the representation of administrative interests and resource considerations impede the realization of neoclassical ideals. One example is the development of prison populations in the Minnesota case. The Minnesota commission decided that the size of the prison population should be kept slightly below the available prison capacity. This is perceived as a way to limit costs and excessive punishment while leaving some leeway for occasionally higher demands. As with other goals, this worked in the first year of guideline initiation (93% of prison capacity filled in 1981), but not in three of the four following years (up to 100% in 1982, 104% in 1983, and 108% in 1985). Prison population increased beyond the capacity level following a new determinate-sentence law for drug offenses. The commission could have neutralized the effects of this law on the use of prison capacity through, for example, lowering the dispositional line. The representation of political-administrative interests in the commission, however, prevented such a decision. The state prison administrator, fighting for funds to open a newly built prison argued—successfully—with the prison overcrowding that followed the determinate-sentence law. Again, sentencing is not an isolated event. Sentencing concerns not just neoclassical ideas, but also the interests of numerous actors who can either influence selectivities that precede sentencing (e.g., prosecutors), or adapt sentencing outcomes to their interests by defining the conditions of sentencing (e.g., the administrator in the Minnesota commission).

IV. DISCUSSION AND CONCLUSIONS: WHY NEOCLASSICAL LAW DOES NOT FIT

Sentencing guidelines are a neoclassical strategy of the American criminal justice system to heal several dilemmas of substantivized criminal law. In this article I argue that neoclassical strategies meet serious structural

[16] The findings for Minnesota are, with modifications, confirmed by recently published preliminary, and partly contradictory, results for the federal sentencing guidelines. Nagel (1990, pp. 935–38) describes preliminary observations of charge, fact, and factor bargaining after the implementation of the federal guidelines. Contributions to Champion (1989) point out that federal prosecutors have even more leeway than state prosecutors (pp. 11–14) and that charge bargaining and fact bargaining will increase under the guidelines (p. 132); Champion predicts a "hydraulic effect" that will shift the need to consider limited prison capacities—i.e., substantivized administrative criteria—from the judge to the prosecutor (1989, p. 233; for opposing conclusions see Steury [1989, p. 108]).

and cultural impediments in modern society. My empirical cases and previous research support four hypotheses.

Hypothesis 1

Basic structural and cultural features of modern society that contributed to the substantivation of criminal law continue to be effective and thus hamper neoclassical reformalization in the legislative process. Such societal features include organized working classes and minority populations, the institutionalization of social scientific knowledge, the state's dependence on large corporations, the limits of markets and their dependence on the production of public goods, the existence of lobby groups, and, finally, the interests of legal clients and the legal profession.

Minority and working-class interests in welfare-oriented substantivation were expressed in the construction of the U.S. Sentencing Guidelines through representatives in Congress with working-class and black constituencies. Thus, it is unlikely that the Sentencing Reform Act would have passed Congress without including the welfare-oriented principles of rehabilitation and treatment. Legislation and guideline promulgation were characterized by intense representation of a diversity of lobby groups. The weight of corporate lobbying became obvious when the second draft of corporate guidelines, based on a retributive rationale, was withdrawn. The diversity of value orientations in modern society and their representation in guideline construction made a systematic and logical, desert-based ranking of offenses impossible. The U.S. Sentencing Commission built its guidelines on past sentencing practice, making the average of judges' previous evaluations and rationales the baseline of future evaluations.

The neoclassical assumption that unequivocal and "truthful" programs could emerge from a broad criminal justice perspective proved to be false. Conflicts between different purposes of punishment for the federal level resulted in a compromise that lists all purposes and leaves the choice to implementing actors. This is not surprising in a complex society. In addition, even the actual creation of guidelines, in Minnesota and on the federal level, limited as their neoclassical quality may be, is difficult to achieve. In other cases the realization of guidelines failed altogether for political reasons, such as interest conflicts between different groups, either in the commission (Maine) or in the legislature (New York).

Hypothesis 2

The organizational and occupational effects of substantivation partly hamper the realization of legislative reformalization. They include the

professionalization and internalized substantive rationales of criminal justice personnel, and substantive administrative demands. Particularly noteworthy is the judges' long-standing interest in maintaining a high degree of autonomy in judicial decision making. It is then not surprising that the influence of judges in the federal case was visible in the legislative process, in the commission, and during the promulgation of the guidelines in establishing relatively generous departure options (Nagel 1990, p. 938; see also Galegher and Carroll 1986, pp. 233, 236).

Social science expertise produced seemingly paradoxical effects. While the emergence of social science is one of the motors of substantivation (see also Stryker [1989, 1990a] on technocratization), social scientists in this policy-making process contributed to a program of reformalization. They did so by working within a set policy frame and disregarding major sociological insights that reach beyond this frame.

Hypothesis 3

The attempt to reformalize criminal law is also problematic due to the strategy chosen. The supporters of neoclassicism suggest that reformalization, that is, the recreation of sociologically formal rationality, should be pursued through the creation of coherent, logical, and gapless instruments of central guidance, that is to say, by increasing the logically formal rationality in legal thought. Here the neoclassical strategy is based on the idea of a strong central state able to create unequivocal and "truthful" programs and to implement these programs through clear bureaucratic organizations. This idea, however, has become questionable in all modern societies and continues to be so in the United States.

The failure of policymakers to guarantee exact implementation became obvious in Minnesota, where the reduction of sentencing purposes and the association of offenses to the seriousness scale was not always accepted by the judges. They reintroduced the purpose of treatment, neglected by the commission, and clearly reduced the prison term proposed by the guidelines for sexual offenders. The commission tolerated their departures, but kept the proposed prison term high for political and general deterrence reasons. It sacrificed the principle of truthfulness while constructing fiction in cooperation with the judges. Although fictions are a necessary element of law (Scheppele 1988), they are contradictory to the neoclassical philosophy and may spoil its instruments. In addition, the continuing exposure of fictions by social scientists is more likely to hurt criminal law than other types of law since criminal law's legitimacy—given the seriousness and forcefulness of its interventions—depends in particular on the criteria of justice and equal treatment. Similarly, the idea of stressing "just desert" was strongly moderated in the

Minnesota case by giving considerable weight to the incapacitation prin-
ciple favored by criminal justice actors, whose views were informed by
administrative and public safety rationales.

Hypothesis 4

*Those elements of the original neoclassical program of reformalization that
are successfully built into the policy instrument are endangered during the
implementation process.* This again arises from the organizational effects
of substantivation. The implicit neoclassical assumption of clear imple-
mentation structures proved to be false. I have already pointed to the
decline of the bureaucratic organization and the rise of professional spe-
cialists. As a result, bargaining and negotiation, instead of rule by hierar-
chy, became typical characteristics of modern organizations, and these
emerging behaviors were themselves opposed to the deductive applica-
tion of sentencing norms through guidelines.

The implementation structure of the courts certainly cannot be altered
by a sentencing commission according to its neoclassical program. Judges
gain their positions through political procedures: they are either nomi-
nated and confirmed or they are elected. Depending on their constituency
and background they are more or less dedicated to the fulfillment of the
substantive rationales (Levin 1977) that are disregarded by sentencing
guidelines. This lack of identification with the rationale of guidelines in
combination with the functional and organizational interconnection of
criminal law with other policy areas such as welfare, health, economic,
and fiscal, leaves little chance that judges will straightforwardly imple-
ment the guidelines. The Minnesota case offered instructive examples for
this thesis.

Galegher and Carroll (1986) stress that successful program innovation
is less likely to succeed the more complex and decentralized the imple-
mentation structure is. In accordance with my observations, they state
that "sentence determinations are largely a product of . . . negotiations
among members of . . . various occupational specialties. . . . Thus, at
least with respect to sentencing decisions, the court may be described as
a complex, decentralized organization" (Galegher and Carroll 1986, p.
239). Although determinate sentencing guidelines intend to reduce bar-
gaining, it is likely that much sentence bargaining will be transformed
into charge and fact bargaining. The power of the prosecutor will in-
crease. Disparities in charging decisions will be invisible.

The findings of my case study are confirmed for mandated penalties
in Massachusetts, for California's sentencing scheme (Feeley 1983, pp.
148, 154–55), and for the "ban" on plea bargaining in felony cases in
California. Here bargaining was shifted from the forbidden sphere into

earlier and less visible stages of the criminal justice process (McCoy and Tillman 1986, pp. 50, 65). These cases confirm our thesis that the impact of any law in the criminal justice system is bound by the internal organizational dynamics and complexity of those decision-making processes in which sentencing is embedded. Warnings would have been available in the literature on the sociology of organizations, particularly the concept of institutionalized organizations (Meyer and Rowan 1977) and its application to criminal justice as a loosely coupled organizational system (Hagan, Hewitt, and Alwin 1979). Studies that do not control for these conditions (e.g., Klein, Petersilia, and Turner [1990] on the California guidelines) give highly problematic guidance to policymakers.

While I found support for all four hypotheses, the neoclassical movement also had successes and certainly major—even though often unintended—consequences in the United States. First, the spread of neoclassical *ideas* is impressive. Second, these ideas initiated numerous policy innovations. Third, some neoclassical ideas influenced the substance of policy instruments. Fourth, impacts can also be registered on the implementation level, for example, in an increase in imprisonment (however, on resource limits see Pontell [1984]) and a transfer of disparities from sentencing to charging. We have also seen that neoclassicism has relatively good chances in specific phases, such as the selection of commissioners, characterized by particular and exceptional conditions: little publicity, stress on technocratic expertise, a small group of actors with relative ideological agreement with regard to neoclassical principles, and the absence of administrative and fiscal constraints. Potential reasons for the relative chances of neoclassicism in the United States were discussed. They include the pluralist structure of interest intermediation, the personalistic univeralism of the United States political system, and the strong cultural sedimentation of ideas of individual rights and responsibilities.

These modifications notwithstanding, my research and previous experiences support my hypotheses and demonstrate the limits of neoclassicism. Guidelines do not appear to be the remedy they were meant to be. Often their construction fails, and even positively evaluated guidelines face enormous implementation problems. Rather than a remedy, guidelines may be a purely symptomatic treatment that allows the disease—the problematic consequences of substantivized law—to spread. The visibility of disparities in criminal justice may be reduced as they are shifted from sentencing into other areas such as prosecution. In this case guidelines would be a very dangerous treatment.

I have documented how, in Max Weber's terms, formal-rational law of modernizing societies gave way to law of the interventionist state, which is characterized by substantive rationality. These developments reflect the formation and articulation of class and other special group

interests in modern states, the growth of the external effects of private production and the resulting need for the production of public goods, changing perceptions of social conditions, professionalization, the growth and internal differentiation of public administration, and related shifts from conditional to purposive state programs. Neoclassical instruments—for example, sentencing guidelines—do not respect this general development; they are thus anachronistic. The transformation of neoclassical *philosophy* into social practice is limited by the *sociologic* of modern societies. Sociological reason suggests other directions of reform.[17]

REFERENCES

Abel, Richard, ed. 1982. *The Politics of Informal Justice* New York: Academic Press.
Alber, Jens. 1982 *Von Armenhaus zum Wohlfahrtsstaat.* Frankfurt Campus.
Benson, Kenneth J. 1975. "The Interorganizational Network as a Political Economy " *Administrative Science Quarterly* 20.229–49.
Benson, Michael, and Esteban Walker. 1988 "Sentencing the White Collar Offender." *American Sociological Review* 53:294–302
Bentham, Jeremy (1780) 1948. *An Introduction to the Principles of Morals and Legislation.* New York Hafner.
Black, Donald. 1976 *The Behavior of Law* Orlando. Academic Press.
———. 1987. "Compensation and the Social Structure of Misfortune " *Law and Society Review* 21:563–84.
——— 1989. *Sociological Justice.* Cambridge. Cambridge University Press.
Champion, Dean J., ed. 1989. *The U S. Sentencing Guidelines: Implications for Criminal Justice* New York. Praeger.
Cullen, Francis T., William J. Maakestaad, and Gray Cavender 1987. *Corporate Crime under Attack: The Ford Pinto Case and Beyond.* Cincinnati: Anderson.
Durkheim, Émile 1893. *De la division du travail social* Paris: Alcan
——— 1899/1900. "Deux lois de l'évolution pénale." *Année Sociologique* 4:65–95.
Edelhertz, Herbert, and Gilbert Geis. 1974. *Public Compensation to Victims of Crime.* New York Praeger.
Eder, Klaus. 1986. "Eine Theorie der modernen Strafrechtsentwicklung " *Kultur und Gesellschaft,* special vol of *Kolner Zeitschrift fur Soziologie und Sozialpsychologie* 27·232–62.
Ewing, Sally. 1987 "Formal Justice and the Spirit of Capitalism Max Weber's Sociology of Law." *Law and Society Review* 21.487–512.
Feeley, Malcolm M. 1983. *Court Reform on Trial: Why Simple Solutions Fail* New York. Basic Books.
Fine, Gary Alan. 1984. "Negotiated Orders and Organizational Cultures." *Annual Review of Sociology* 10·239–62.
Foucault, Michel. 1975 *Surveiller et punir: La naissance de la prison.* Paris: Edition Gallimard. (In English. 1979. *Discipline and Punish· The Birth of the Prison.* New York: Random House).
Frehsee, Detlev 1987. *Schadenswiedergutmachung als Instrument strafrechtlicher Sozialkontrolle.* Berlin· Duncker & Humblot.
Friedman, Lawrence M. 1985. *A History of American Law.* New York: Simon & Schuster.

[17] See, e.g., Black (1976, p. 137; 1987; 1989, pp. 41 ff.).

Galegher, Jolene, and John S. Carroll. 1986. "Thinking Things Through: A Concep-
tual Analysis of the Construction and Implementation of Sentencing Guidelines."
Pp. 226–48 in *The Impact of Social Psychology on Procedural Justice*, edited by
Martin F Kaplan Springfield, Ill.: Thomas

Garland, David. 1985. "The Criminal and His Sentence: A Critical Account of the
Formation of Criminology at the End of the Nineteenth Century." *British Journal
of Criminology* 25·109–37.

Gewirth, Alan. 1978. *Reason and Morality* Chicago. University of Chicago Press.

Ginsburgh, Norman. 1979. *Class, Capital, and Social Policy*. London: Macmillan

Greenberg, David F , and Drew Humphries. 1980. "The Cooptation of Fixed Sen-
tencing Reform." *Crime and Delinquency* 26·206–25.

Grundgesetz. (1948) 1974. Munich. C. H. Beck

Habermas, Jurgen. 1979. *Toward a Rational Society.* Boston· Beacon.

———. 1975. *Legitimation Crisis*. Boston. Beacon.

Hagan, John, and Patricia Parker. 1985. "White Collar Crime and Punishment."
American Sociological Review 50:302–16

Hagan, John, John D. Hewitt, and Duane F. Alwin. 19' 9. "Ceremonial Justice:
Crime and Punishment in a Loosely Coupled System." *Social Forces* 58:506–27.

Hurst, James Willard. 1956. *Law and the Conditions of Freedom in the Nineteenth
Century United States* Madison· University of Wisconsin Press.

Kant, Immanuel 1965 *Metaphysical Elements of Justice*. Indianapolis: Bobbs-
Merrill.

Katzmann, Robert A. 1980. *Regulatory Bureaucracy: The Federal Trade Commission
and Antitrust Policy*. Cambridge, Mass. MIT Press.

Kelman, Steven. 1981. *Regulating America, Regulating Sweden: A Comparative Study
of Occupational Safety and Health Policy* Cambridge, Mass.: MIT Press.

Klein, Stephen, Joan Petersilia, and Susan Turner. 1990. "Race and Sentencing Deci-
sions in California." *Science* 247:812–16.

Knapp, Kay A 1987. "Implementation of the Minnesota Guidelines· Can the Innova-
tive Spirit Be Preserved?" Pp 127–41 in *The Sentencing Commission and Its
Guidelines* by Andrew von Hirsch, K A. Knapp, and M. Tonry. Boston: North-
eastern University Press.

Landes, William M., and Richard A. Posner. 1987. *The Economic Structure of Tort
Law*. Cambridge, Mass.: Harvard University Press.

Levin, Martin A 1977. *Urban Politics and the Criminal Courts* Chicago· University
of Chicago Press

Luhmann, Niklas. 1976. *Funktionen und Folgen formaler Organisation*. Berlin
Duncker & Humblot.

———. 1985 *A Sociological Theory of Law*. London. Routledge & Kegan Paul.

Martinson, R. 1974 "What Works? Questions and Answers about Prison Reform "
Public Interest 12:22–55.

Marx, Karl, and Friedrich Engels. 1938 *The German Ideology* London: Lawrence
& Wishart.

McCoy, Candace, and Robert Tillman. 1986. *Controlling Felony Bargaining in Cali-
fornia: The Impact of the "Victims' Bill of Rights."* Sacramento· California Depart-
ment of Justice.

Meyer, John W., and Brian Rowan. 1977 "Institutional Organizations: Formal
Structure as Myth and Ceremony." *American Journal of Sociology* 83:340–63.

Murphy, Jeffrie G., and Jules L. Coleman. 1990 *Philosophy of Law: An Introduction
to Jurisprudence*. Boulder, Colo.: Westview Press.

Nagel, Ilene H. 1983. "The Legal/Extra-Legal Controversy. Judicial Decisions in
Pretrial Release " *Law and Society Review* 17:481–515.

——— 1990. "Foreword: Structuring Sentencing Discretion: The New Federal Sen-
tencing Guidelines." *Journal of Criminal Law and Criminology* 80:883–943.

Nonet, Philippe, and Philip Selznick. 1978 *Law and Society in Transition: Toward Responsive Law* New York: Octagon.

Olson, Mancur, Jr. 1965 *The Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, Mass . Harvard University Press.

Pontell, Henry N. 1984. *A Capacity to Punish: The Ecology of Crime and Punishment*. Bloomington: Indiana University Press.

Pontell, Henry, Wayne Welsh, Patrick Kinkade, and Matthew Leone. 1988. "Incarceration as a Deviant Form of Social Control: Jail Crowding in California." Paper presented at the meeting of the Pacific Criminological Association, Las Vegas.

President's Commission on Law Enforcement and the Administration of Criminal Justice. 1967. *The Challenge of Crime in a Free Society*. Washington, D.C . Government Printing Office

Robinson, Paul H. 1987. *Dissenting View of Commissioner Paul H. Robinson on the Promulgation of Sentencing Guidelines by the U.S. Sentencing Commission*. Washington, D.C.: Government Printing Office.

Rosner, Richard, and Ronnie B. Harmon, eds. 1989 *Criminal Court Consultation*. New York: Plenum.

Roth, Günther. 1987. *Politische Herrschaft und Persönliche Freiheit: Heidelberger Max Weber-Vorlesungen*. Frankfurt· Suhrkamp.

Samaha, Joel. 1988. *Criminal Justice* St. Paul, Minn.: West.

Savelsberg, Joachim J. 1987. "The Making of Criminal Law Norms in Welfare States: Economic Crime in West Germany." *Law and Society Review* 21:529–61.

———. 1988. "Rationalities and Experts in the Making of Criminal Law against Economic Crime." *Law and Policy* 10:215–52

Savelsberg, Joachim J., and Peter Bruhl. 1988. *Politik und Wirtschaftsstrafrecht: Eine soziologische Analyse von Kommunikationen, Rationalitaten, und Macht*. Leverkusen Leske & Budrich.

Scheppele, Kim Lane. 1988. "Jeremy Bentham and the Functions of Legal Fictions." Paper presented at the 24th annual meeting of the Law and Society Association.

Schluchter, Wolfgang 1981. *The Rise of Western Rationalism: Max Weber's Developmental History*. Berkeley· University of California Press.

Schmitter, Philippe C. 1982. "Reflections on Where the Theory of Neo-corporatism Has Gone and Where the Praxis of Neo-corporatism May Be Going." Pp. 259–79 in *Patterns of Corporatist Policy-Making*, edited by Gerhard Lehmbruch and Philippe C Schmitter Beverly Hills, Calif.: Sage.

Shavell, Steven. 1987. *Economic Analysis of Accident Law* Cambridge, Mass.: Harvard University Press

Steury, Ellen H 1989. "Prosecutorial and Judicial Discretion." Pp 93–110 in *The U.S. Sentencing Guidelines: Implications for Criminal Justice*, edited by Dean J. Champion. New York: Praeger.

Strafprozessordnung. n.d. Bonn: Deutscher Bundestag.

Strauss, Anselm. 1978. *Negotiations: Varieties, Contexts, and Social Orders*. San Francisco· Jossey-Bass.

Stryker, Robin. 1989. "Limits on Technocratization of Law." *American Sociological Review* 54:341–58.

———. 1990*a*. "A Tale of Two Agencies: Class, Political-Institutional, and Organizational Factors Affecting State Reliance on Social Science " *Politics and Society* 18:101–41.

———. 1990*b*. "Science, Class, and the Welfare State. A Class-centered Functional Approach " *American Journal of Sociology* 96:684–726.

Teubner, Gunther. 1987. "Juridification: Concepts, Aspects, Limits, Solutions." Pp. 3–48 in *Juridification of Social Spheres: A Comparative Analysis in the Areas of Labor, Corporate, Antitrust and Social Welfare Law*, edited by Gunther Teubner. Berlin/New York: Walter de Gruyter.

Therborn, Göran 1978 *What Does the Ruling Class Do When It Rules? State Appara-tuses and State Power under Feudalism, Capitalism, and Socialism.* London: New Left Books.

Turkel, Gerald. 1980. "Rational Law and Boundary Maintenance: Legitimating the 1971 Lockheed Loan Guarantee." *Law and Society Review* 15:41–77.

Unger, Roberto Mangabeira. 1976. *Law in Modern Society: Toward a Criticism of Social Theory.* New York: Free Press.

U S. House of Representatives, Subcommittee on Criminal Justice. 1988. *Hearings on the Sentencing Guidelines.* Washington, D.C.: Government Printing Office.

U.S. Senate, Committee on the Judiciary 1977. *Reform of the Federal Criminal Laws,* pt. 13. Hearings before the Subcommittee on Criminal Laws and Procedures of the Committee on the Judiciary, U S Senate. Washington, D.C.. Government Printing Office.

————. 1979*a Reform of the Federal Criminal Laws,* pt 14. Hearings before the Committee on the Judiciary, U.S. Senate Washington, D.C.: Government Printing Office.

————. 1979*b. Reform of the Federal Criminal Laws,* pt. 15. Hearings before the Committee on the Judiciary, U.S. Senate. Washington, D.C.: Government Printing Office.

————. 1981. *Reform of the Federal Criminal Laws,* pt. 16. Hearings before the Committee on the Judiciary, U.S. Senate. Washington, D.C.: Government Printing Office.

————. 1983. *Comprehensive Crime Control Act of 1983: Hearings before the Subcom-mitee on Criminal Law.* Washington, D.C.: Government Printing Office.

————. 1985. *Confirmation Hearings for the U.S. Sentencing Commission.* Washing-ton, D.C.: Government Printing Office.

U.S. Sentencing Commission. 1987. *Sentencing Guidelines and Policy Statements* (April 13) Washington, D.C.: Government Printing Office.

Von Hirsch, Andrew. 1987*a.* "Structure and Rationale: Minnesota's Critical Choices." Pp. 84–106 in *The Sentencing Commission and Its Guidelines,* by A. von Hirsch, K. A Knapp, and M Tonry Boston: Northeastern University Press.

————. 1987*b.* "The Sentencing Commission's Function." Pp. 3–15 in *The Sentenc-ing Commission and Its Guidelines,* by A. von Hirsch, K. A. Knapp, and M. Tonry Boston: Northeastern University Press.

Von Hirsch, Andrew, Kay A. Knapp, and Michael Tonry. 1987. *The Sentencing Commission and Its Guidelines* Boston: Northeastern University Press.

Washington Post. 1990*a.* "Justice Department Shifts on Corporate Sentencing." April 28.

———— 1990*b.* "No. 2 Aide at Justice Quits Post." May 12.

Weber, Max. 1968. *Economy and Society.* Berkeley: University of California Press.

Weir, Margaret, and Theda Skocpol. 1985. "State Structures and Possibilities for 'Keynesian' Responses to the Great Depression in Sweden, Britain, and the United States." Pp. 107–63 in *Bringing the State Back In,* edited by Peter B. Evans, Dietrich Rueschemeyer, and Theda Skocpol Cambridge· Cambridge University Press.

Wheeler, Stanton, Kenneth Mann, and Austin Sarat. 1988. *Sitting in Judgement: The Sentencing of White Collar Criminals.* New Haven, Conn: Yale University Press.

Wheeler, Stanton, David Weisburd, and Nancy Bode. 1982. "Sentencing the White Collar Offender: Rhetoric and Reality." *American Sociological Review* 47·641–59.

Wilson, James Q. 1975. *Thinking about Crime* New York: Basic Books.

A Notable Administration: English State Formation and the Rise of Capitalism[1]

Derek Sayer
University of Alberta

England is, according to both Marx and Weber, the classic ground of modern rational capitalism. Yet England's political history and institutions strikingly deviate from what Weberian or Marxist ideal-types of capitalist development might lead us to expect. This article argues that these deviations are important in explaining why England became the home of capitalism in the first place. Particular stress is put upon the earliness of England's formation as a national state, or the continuities of its major legal and political institutions, and what are often seen as their amateurish "irrationalities" in molding, over the very *longue durée*, a civil society in which capitalist economy was possible. If so, the pertinence of Marxist and Weberian ideal-types to a *historical* sociology of capitalism, whether in England or elsewhere, needs very seriously to be reconsidered.

I

In some jottings of 1892, Friedrich Engels confessed to finding the "logical incongruities" of the English state "a sore trial to the reasoning mind" (Engels 1962, p. 529). One can sympathize with his predicament. England presents a peculiar paradox for social theory. It is by common consent "the home of capitalism" (Weber [1920–23] 1966, p. 251): its "classic ground" (Marx [1867] 1967, p. 8). Yet England's political institutions and culture have not taken the forms that the Marxist and Weberian ideal-types of the modern, bourgeois polity might lead us to expect. The discrepancies are many and serious. My argument in this article is that these logical incongruities form an important part of the explanation of why England became the home of capitalism in the first place. If

so, the pertinence of these ideal-types to a historical sociology of the origins of capitalism seriously needs to be rethought.

One point should perhaps be clarified at the start. It is not Marx and Weber, so much as what others have abstracted from them, that I am criticizing here. Marx was a far more cautious generalizer, and a more scrupulous historian, than many have acknowledged (Sayer 1987, 1989). Weber made it quite clear that his ideal-types were neither descriptions, nor averages, nor hypotheses, but analytical constructs that one-sidedly emphasize those features of a phenomenon most relevant to the interests of the investigator (Weber [1904–18] 1949, pp. 89 ff.; [1904–5] 1974, pp. 13–31). He spells out the point of view for his main interest in rational capitalism in his 1920 introduction to *The Protestant Ethic*. He sees it as part of an inquiry into the rationalization that he thinks distinguishes the historical trajectory of occidental civilization from others, and underpins what is most distinctive to modern life (Lowith 1982; Sayer 1991). Marx (1967, p. 10; [1857–58] 1973, pp. 460–61), too, was much more concerned to "lay bare the economic law of motion of modern society" than to write "the real history of the relations of production." Within *these* contexts of inquiry, Marxist and Weberian ideal-types have undoubted value. But it is exactly this present-centeredness that makes them of dubious utility to a *historical* sociology. Weber himself was emphatic that "early modern capitalism did not *originate* in bureaucratic model states where bureaucracy was a product of the state's rationalism," and even "advanced capitalism . . . was not at first limited to these countries, in fact, not even primarily located in them." He made the very different claim that "*today,* capitalism and bureaucracy have found one another and belong intimately together" (Weber 1978, p. 1465; emphasis added). Others have not always been so mindful of this distinction.

II

For Weberians the bourgeois ethos is typically urban, and Weber himself lays stress, in his more general pronouncements, on the role of the free medieval European city in its formation. Cities are "the seat of commerce and industry," requiring "a continuous provision of the means of subsistence from without." They are also critical agencies of a general cultural rationalization, whether in science, religion, art, or politics. The city as a political community that possesses its own laws and elected administrative officials is a uniquely Western institution, as is the concept of citizenship. The basis for the latter, Weber maintains, resides in "the *coniuratio,* the brotherhood in arms for mutual aid and protection," something absent from the Orient because "the army of the prince was older than

the city." The medieval European city, unlike its ancient predecessors, was also for Weber the source of two eminently bourgeois virtues, the "equalization of classes and removal of unfreedom," which have their source in the circumstance that here it is the status of a householder, not that of a landowner, that confers citizenship. In Northern Europe, he says, this was reinforced by the frequent exclusion of nobles and high officials from town residency (Weber 1966, chap. 28).

Weber is not alone in seeking the origin of bourgeois mores far from the "idiocy of rural life" (Marx, in Marx and Engels [1847–48] 1976b, p. 488). Chirot (1985) focuses centrally on the distinctiveness of the European city in his recent account of "the rise of the West." Plenty of Marxists do so, too, if for different reasons. Paul Sweezy (1978) follows Henri Pirenne (1974) in seeing the solvent of feudalism as the urban production stimulated by the 11th-century revival of long-distance trade, while Perry Anderson (1974a, p. 205) credits "the *urban* sector, structurally sheltered by the parcelization of sovereignty in the medieval polity," with "decisively alter[ing] the outcome of the class struggle in the rural sector" after the Black Death. It was, he claims, the "radiation of market relationships into the surrounding countryside," and—a very dubious assertion, ably challenged by Robert Brenner (1985)—burgesses' assistance to peasant revolts, that opened up the high road to agrarian capitalism (Anderson 1974a, p. 205). For Marx, too, Asiatic cities contrasted with those of medieval Europe in being mere "royal camps" (Marx 1973, p. 479).

Strange, then, that one of England's major differences with much of continental Europe, during the Middle Ages and after, was the comparative absence of political autonomy of urban communes; and that this was bound up with "the extraordinary concentration of political power in England after William the Conqueror" (Weber 1966, p. 246). Here the army of the prince (or more precisely the strong central capability that it established) did preempt the development of the free city. By 1700 England's urban profile contrasted doubly, and very strikingly, with much of continental Europe's. The domination of London, the administrative capital, then by far the largest conurbation in Europe with some 500,000 souls, had no Continental parallel, but the three next-largest towns in the realm, Norwich, Bristol, and York, housed a paltry 25,000–30,000 people each. England boasted no major provincial centers comparable to Marseilles, Lyon, or the great cities of Germany and Italy. Weber himself remarks that "the great modern commercial and industrial cities of England arose outside the precincts . . . of the old privileged corporations" (1978, p. 1331).

Nor was there (as Weber [1966, p. 246] also acknowledges) any exclusion of the nobility from urban affairs. On the contrary, it was the landed

gentry who most often provided the boroughs with their Members of Parliament (M.P.'s) from the Tudor period to the 19th century, something that needs to be borne in mind in the context of Weber's claim that cities in England, though lacking "separate and individual political ambitions . . . defended their interests in Parliament as a group" (Weber [1910–20] 1978, p. 1324). The lack of polarization of aristocracy and urban patriciate is evident in their intermarriages. According to Laslett (1971, pp. 49–51), there is little familial continuity to be found within city oligarchies in England, by comparison with either Italy or Holland. He explains this in terms of the ease of marriage between gentle and merchant families, and gentlemen's own participation in urban economic pursuits. Weber himself says—a point very relevant to the thesis of this paper—that this fusion of urban and rural elites "in the type of the *gentleman* was greatly facilitated by their common ties to the office of the justice of the peace" (Weber 1978, p. 1060).

This relates to a wider English peculiarity. By 1700 or so, the gentlemen and aristocrats who (then and for long after) ran English governance were in many ways quite clearly a bourgeois class. They farmed for profit, speculated, invested in colonial ventures, exploited coal and iron deposits on their lands and, as E. P. Thompson once put it, "valued each other not in the scales of breed and antiquity but in round annual sums" (1978, p. 44). Yet this "superbly successful and self-confident capitalist class" also lived on country estates, rode to the hunt, and administered their counties as Justices of the Peace, Deputy Lieutenants, and Overseers of the Poor in ways that were abidingly patrimonial and deeply rooted in tradition (Thompson 1978, p. 43). This is but one instance of a broader phenomenon, which again sits oddly with Weberian ideal-types. Well into the 20th century, despite waves of attempted rationalization (like the "revolution in government" of the 1830s), a great deal in the English polity has remained conspicuously *un* "modernized."[2] A small but revealing example is the custom of awarding titles of honor that derive from the status systems of "the past," something deliberately done away with in France and the United States during their revolutions. Of more consequence, perhaps, is the enduring importance of the Privy

---

[2] I have used quote marks here and elsewhere in the paper in order to signal that English historical experience should lead us to question the routine categories of much sociology, particularly those which hinge upon some variant of a tradition/modernity dichotomy. The Duke of Westminster, who is the wealthiest individual in England after the Queen, is a *modern* figure, not a feudal relic. Conversely, to describe medieval English state formation as precocious implies that state formation is somehow an improper aspect of premodern societies. It is exactly the ideal-types I am criticizing that sustain such descriptions, and the related perception of England's history as "peculiar."

Council and royal prerogative as enabling and legitimating powers for the modern executive.

For much of its long history, and certainly for long after capitalist forms of economy were well-entrenched, England had neither a formally rational law nor a rational bureaucracy, in Weber's sense of these terms. England did exhibit what, by Continental standards, was a quite remarkable precocity in the development of national agencies of adminstration and justice. But, in their forms, these did not fit rational-bureaucratic ideal-types, and their lack of fit has been what is most enduring and characteristic about them. The Common Law did not operate on the basis of "calculable rules," and has never been codified. Weber himself says that such codification would have been "practically impossible" (1978, p. 890). It is a peculiar amalgam of Parliamentary statute and judicial precedent, whose most evident feature is what Weber calls its "practical adaptability" (1978, p. 890). For centuries its implementation was largely in the hands of unpaid and legally untrained Justices of the Peace (J.P.'s), whose jurisdiction Weber terms "patriarchal, summary and highly irrational" (1978, p. 891). Common Law could scarcely "be counted upon, like a machine," as Weber's *General Economic History* says rational capitalism requires (1966, p. 252), although in *Economy and Society,* when discussing the Anglo-American legal tradition, he is more cautious, concluding that "the legal systems under which capitalism has been prospering differ profoundly from each other even in their ultimate principles of formal structure" (Weber 1978, p. 890). Nor was it much influenced by the revival of Roman law that engulfed much of the Continent, as both Weber (1978, p. 976; 1974, p. 77) and Marx (in Marx and Engels [1845–46] 1976a, p. 91) remark. In Weber's own words, "the first and most highly developed capitalist country . . . retained a less rational and less bureaucratic judicature" (1978, p. 977) than its Continental rivals. England's *industrial* cities "frequently displayed a completely archaic judicial structure" of courts baron and leet (Weber 1978, p. 1331).

Even if there were arenas of state, like the 18th-century Excise, where a case might be made (Brewer 1989) for the existence of a "Weberian" bureaucracy before the 19th century, bureaucracy was not the usual vehicle of administration in England. In the single French province of Normandy, by the 1630s, there were more than 3,000 salaried, professional, and full-time state officials. The Kingdom of Castile boasted 530 major posts, and above 30,000 lesser places under the Exchequer alone. Venal as these officeholders may often have been, the comparable figure for the whole of England during the same decade was at most "some hundreds": a telling contrast (Aylmer 1974, pp. 440, 452). As late as 1871, Karl Marx could fume against English "self-government" by "the holders of broad

acres, long purses, and empty heads" (in Marx and Engels [1871] 1986, p. 464). Sixteen years earlier he had ridiculed "the British Constitution" as an "antiquated, obsolete, out of date compromise" whose "mask" would imminently be dropped, to reveal "the real features of Britain's political physiognomy" (in Marx and Engels [1855] 1980, pp. 53, 56). In the event, bureaucratized forms of governance, like the new system of policing, spread slowly and patchily through the counties, even during the later 19th century (Steedman 1984). Weber knew this, too. English administration, he says, was "an administration of notables," and "England . . . the slowest of all countries to succumb to bureaucratization" (1970, p. 228).

Pervasive in all this is an antique, gentlemanly amateurism, precisely the ethos of all-round cultivation that Weber sees as antithetical to the specialized officialdom that is required and spread by capitalism (Weber 1978, pp. 998–1002). It is, I think, very revealing that that biographer of self-made men, Samuel Smiles, should have entitled the concluding chapter of his *Self-Help,* a paean to Victorian values that sold over a quarter of a million copies between its publication in 1859 and Smiles's death in 1904, "Character: The True Gentleman" (Smiles [1859] 1986). Again, Weber was well aware of this gentlemanly ethic, though whether it was quite so "rationalized" under the impact of Puritanism as he suggests (1978, p. 1063) is debatable. With regard to Puritanism, England is again very far from ideal-typical. Weber is clear that Protestantism is *not* indispensable to capitalism (1974, p. 91). But the form it assumed in the classic ground might nevertheless give pause for thought, given what is conventionally taken to be the Weber thesis. Richard Baxter, who looms so large as a Puritan exemplar in *The Protestant Ethic,* was a nonconformist who suffered imprisonment and public burning of his books after the Restoration. The Reformation in England was an act of state, motivated much more by considerations of political expediency than religious fervor; and although England was indeed Protestant thereafter, it was not the Puritans who were the final victors in the Civil War. Between the Restoration and 1828, dissenters found themselves disadvantaged politically, if less so than papists. Those qualities of ascetic Protestantism, from the doctrine of predestination (Weber 1974), to sectarian forms of discipline (Weber 1970, pp. 303–22), which Weber regards as having had an elective affinity with the spirit of capitalism, were largely missing from the Anglican Church to which most English men and women, from the 16th to 19th centuries, were constrained to belong. So were the individualism and the antiauthoritarianism stressed by writers of Marxist inclination, from Tawney (1938) to Hill (1974). If doctrinally Protestant, the Church of England was organizationally episcopalian, and for centuries numerous ecclesiastical livings were in the gift of

landed gentlemen. To this day the archbishop of Canterbury is appointed by the prime minister. This church was, from the 1530s onward, a state church—and a central ingredient in England's "peculiar" polity.

III

Marxist paradigms do not fare any better in the face of English "peculiarities." Participants in the famous transition debate of the 1950s and 1960s acknowledged the difficulty of applying the Marxist categories of feudal and capitalist at all to the English polity during the two (or three, or four) centuries when it was seemingly neither one thing nor the other (Hilton 1978). This does not affect only the modern end of "the transition," those bourgeois aristocrats and gentlemen of the 17th and 18th centuries; Alan Macfarlane (1986) detects a money economy and a free market in land and labor by, at latest, the 13th century. The class character of the state during any period between the 16th and the 19th centuries is eminently debatable. Furthermore, centralization of state power out of the "checkered . . . anarchy of conflicting medieval powers" (Marx, in Marx and Engels 1986, p. 483) has been seen by many Marxists as a key staging-post on the road to *bürgerliche Gesellschaft*. Marx, in his studies of France, a country he considered to be "classical" (as on occasion did Weber [1978, p. 989]), concluded that absolutism was a crucial "weapon of modern society in its struggle of emancipation from feudalism" (in Marx and Engels 1986, p. 483). But as we shall see, checkered anarchy—or what some later Marxists (e.g., Anderson 1974a, p.148) call parcelization of sovereignty—was not characteristic of medieval English governance, nor can the Tudor/Stuart monarchy plausibly be described as an absolutism.

The difficulty which has most exercised Marxists, however, is to find a "bourgeois revolution" in English political history. The usual candidate for this accolade is the upheavals of 1640 to 1689. The choice has much to be said for it, *if* we accept the implicitly teleological framework within which such judgments are made. There was a rebellion going on during the 1640s and 1650s, and neither the Restoration of 1660 nor the settlement of 1689 merely restored the *status quo ante*. They considerably curbed the English crown, in ways that were undoubtedly conducive to the further advancement of capitalism. The abrogation of feudal tenures helped agricultural "improvement," the Navigation Acts fostered commerce, and curtailment of the royal prerogative removed what might otherwise have become major obstacles to subsequent capitalist development. Common Law and Parliament gained a new supremacy over royal "arbitrariness"; both facilitated England's *embourgeoisement,* leaving men of property enviably free to enjoy their very substantial liberties. It

is also improbable that the development of capitalist relations within agriculture in the decades before 1640 did not play *some* part in inducing the social strains that propelled England toward civil war in the first place (Manning 1978; Hill 1981; Hirst 1981; but, cf. Goldstone [1988], who puts these events in a global demographic perspective).

But a major sociological difficulty remains, and it is quintessentially English. Even Christopher Hill concedes that the Rebellion was bourgeois neither in its social composition nor its objectives, might have been avoided anyway had Charles I been better advised and less obdurate, and, in its ultimate settlement, confirmed the political domination of the pillars of the *ancien régime* (Hill 1980, pp. 110–12, 118, 124)—all points that most modern authorities would endorse. We *can* of course so sever structure from agency that the events of 1640–89 qualify as a bourgeois revolution by dint of their outcomes for capitalism alone. But this does not resolve the problem. Worse, this view of the 1640s from the perspective of the future obscures those social forces and political agendas that actually were in conflict at the time.

Those measures of reform that survived the Restoration were for the most part (abolition of feudal tenures and the Navigation Acts are exceptions) passed at the outset of the conflict, during 1640–42, and it is predominantly these (abolition of High Commission, prerogative courts, patents of monopoly, and non-Parliamentary taxes) that Hill (1980) hails as the essence of his bourgeois revolution. The difficulty is that, as Hexter (1961) observes, all of these were enacted (and with "striking unanimity") by the representatives of England's *traditional* political nation, those gentlemen with large landholdings who made up the overwhelming majority of M.P.'s in the Long Parliament. Attempts (originating with R. H. Tawney [1941*a*, 1941*b*]) to find in this group a protocapitalist "rising gentry" confronting a moribund "feudal" aristocracy have long ago been discredited (Stone 1965). Recent studies suggest, indeed, that before 1640 there was no real opposition party within Parliament, while the independence of the Commons from manipulation by peers has been exaggerated (Russell 1979). Zagorin (1970) found no meaningful correlation between "Country" and "advanced" economic interests.[3] Indeed before 1640 the crown itself was in the van of "progress," being both a large encloser of lands and a backer of improvements like the draining of the fens (Manning 1978), and it *lost* these powers of direct economic intervention in 1660.

---

[3] The terms "Court" and "Country" have been employed by some historians of 17th-century England to identify lines of political cleavage in English society on the eve of the Civil War. There is debate over both the social basis of these groupings and the degree to which either ever formed a coherent faction or lobby. See Stone 1965; Zagorin 1970; Clark 1986.

It is at least arguable that Charles's crime was that in a series of areas ranging from religion to foreign policy, he challenged *established* norms (Clark 1986). What he threatened above all were the powers of the "political nation," powers summed up in the oft-used phrase "self-government at the King's command." Country grievances—ship money, the 11 years rule without Parliament, inflation of honors and the machinery of royal patronage, misuse of prerogative courts, monopolies, and enclosure commissions—were cut from one cloth. In all instances, the entrenched rights of traditional elites to run their own affairs, whether economic, religious, or political, free from arbitrary royal interference were perceived as under attack. "Their own affairs" included the affairs of inferiors. This was a governing class within a polity whose most distinctive feature, from the Middle Ages onward, had been what Marc Bloch called "the collaboration of the well-to-do classes in power" (1967, p. 371).

These rights were indeed amenable to expression in terms of a proud rhetoric of old English liberties. But this was less an anticipation of Tom Paine than a remembrance of that distinctly inegalitarian document, the Magna Charta, which Michael Clanchy (1983, p. 198) calls the first written constitution in Europe. Magna Charta conferred rights only on the free. Hexter (1961) brings this filiation out superbly. The Englishness that was appealed to here had also, by this time, embraced Protestantism. This was not, however, Max Weber's ascetic Puritanism, but Richard Hooker's Anglican compromise that sinewed together church and nation, the Erastian creation of the Henrician and Elizabethan state. Archbishop Laud's policies on doctrine and ritual threatened a long extant ecclesiastical settlement, and the anti-Catholic phobia was (in part at least) an expression of English nationality that the political nation by then considered itself to embody, personally.

More radical, even bourgeois, voices did emerge during the ferment of the Civil Wars, and they may indeed have had important longer-term cultural consequences for England's *embourgeoisement*. Dissent was strongest among "the middling sort." Hill's writings are at their best here, and *The Protestant Ethic* most obviously germane. Those, however, who brought about the only enduring *political* revolutions during the 17th century were England's traditional governors, and they did so in pursuit of what were in many ways highly conservative objectives. What they took issue with in the 1640s was less a "feudal" polity than a disturbingly modernizing king, tainted with suspiciously Continental ideas of Divine Right (and *cujus regio ejus religio*), and they would do so again in 1688–89, taking care, this time, to ensure that the lower orders were not drawn into their quarrels. The English polity was no less the creature of the landed aristocracy in 1689 than it had been in 1450. If anything, in 1660, there was "an actual *strengthening* of the

established order and of the traditional pillars of State, Church, and social hierarchy" (Aylmer 1986, p. 203), with the defeat of religious and political radicalisms. Whatever windfalls the 17th-century revolutions may have brought to capitalism, then, this was far from a conflict in which, as Christopher Hill once put it, "the state power protecting an old order that was essentially feudal was overthrown [and] power passed into the hands of a new class" (Hill 1955, p. 6). It *did,* however, put an end to any royal pretensions toward absolutism, and this, too, is problematic, given the Marxist analyses of the role of the latter, as exemplified by France, in the making of the modern world. England became the workshop of that world under the rule of the political formation that William Cobbett was to christen "Old Corruption" (Thompson 1978, pp. 48–50).

IV

The burden of the argument thus far is simple. Insofar as the rise of capitalism in England was fostered by political factors, these were not such that Marxist or Weberian ideal-types of a modern, bourgeois polity would lead us to expect. Following Weber, the political conditions in which capitalism emerged in its classic ground were not those typically associated with its maturity. We can, of course, more or less persuasively seek out functional equivalents by substituting J.P.s for *Beamten,* improving gentry for urban burghers, and 1689 for 1789. But to do so begs the question of whether it was not precisely the ways in which the English polity differed from these ideal-types that contributed to its becoming the "home of capitalism." In the rest of this article, I shall suggest that this was indeed the case.

To begin by stating the obvious: Great Britain is an island off the Atlantic coast of Europe. Great Britain and England are not of course the same entities, and in what follows I am talking of the English, not the British, state. Britain is an artifact of conquest and colonization, and its unity has been a precarious achievement up to the present day. The geopolitical situation of the English state, however, with its absence of land frontiers that pose a serious military threat, is singular within Europe. This accident of geography has been politically consequential, as many others (most recently, Mann [1986] and Aylmer [1990]) have noted. No internal state formation can be properly comprehended without reference to its external contexts (Chirot 1985; Tilly 1990).

Paul Kennedy's recent major study (1988) of the disastrous economic effects of Continental military ventures for Spain and France suggests to me that England's relative *lack* of Continental involvement, after 1450, may have been significant for its later economic and political develop-

ment. During the critical years for the making of capitalism, the 16th to 18th centuries, England was not squandering productive resources on Continental empire building, nor obliged, to the same degree or in the same ways as Continental powers, to defend itself against others' expansionist predilections. Its state structures were not organized so as to facilitate this objective, with the fetters on private enterprise this certainly would have entailed. The absence in England of a large standing army had direct implications for its state forms, in particular, for the character of the monarchy (as Weber notes, [1978, pp. 970–71]), while England's sea power aided its trade. Defensive reliance on a navy may also have made the landed classes more willing to tax themselves than were their Continental brethren, since the navy did not threaten their power in the way a royal army might. By the 18th century, this capacity to raise taxes—regularly rather than arbitrarily, and at levels that did not kill the goose that laid the golden eggs—became a critical factor in the English government's ability to borrow successfully on international money markets, giving it an edge in commercially lucrative wars. Security of tax revenues also boosted confidence of domestic creditors in the payment of interest, a requisite for stable internal capitalist investment (Brewer 1989).

The fact that England has not been successfully invaded or occupied for the best part of a millenium (a characteristic no other major European state, from France to Russia, shares) may also have been significant. Aylmer (1990) remarks the linguistic unity of England: southern, midland Middle English had established a dominance by the mid-14th century. While not wholly explained by political factors, this was certainly helped by England's geopolitical situation after 1066. In France by contrast, as late as 1863, according to official figures, 8,381 communes out of 37,510 spoke no French (Segal 1988). In the areas that were to become Italy and Germany, linguistic communality reached well beyond political boundaries, a fact that was to have a tragic heritage in the 20th century when the national state had become the norm with ethnicity as its proclaimed basis. But for earlier periods it meant that one major signifier of ethnic identity, linguistic distinctiveness, was unavailable as a plausible element in the forging of a national political community. Linguistic uniformity also allows for (and is in turn promoted by) the standardization of legal and administrative procedures in the vernacular, while linguistic distinctiveness permits the emergence of a consciously patriotic literature. England's freedom from invasion may also have fostered its sense of national uniqueness—whose darker side, harnessed for purposes of empire (and thus commerce) later, is an ideology of national and racial superiority—buttressed its resistance to Continental influences (like Roman law), and (possibly) made the English less inclined to resort to arms

to settle internal disputes than their European neighbors. After the Tudor period, the Civil War apart, most social and political conflicts within England were (comparatively) peacefully managed. The 11 reported dead at Peterloo are small beer when set against the slaughter of 20,000 Paris communards.

But the deeper significance of England's insularity might lie elsewhere. Quite simply, a political form that was, much later, to become general throughout Europe (and the world), was successfully pioneered there rather earlier than elsewhere. England was by no means singular in attempting this; indeed, the templates for its governance were probably Carolingian (Campbell 1986). One of the reasons why a form that collapsed in Europe survived and developed in England, however, may have been the country's political geography. Some frontiers are more defensible than others.

V

Truly national agencies of governance developed much earlier in England than they did elsewhere, or at least they survived where Continental counterparts on which they were probably modeled did not. In Europe Carolingian administration gave way to political feudalism, rather than, as in post-Conquest England, evolving as elements of a centralized polity. Parts of the tapestry of the English state are very old indeed. According to James Campbell (1975, p. 43), the late Anglo-Saxon state exhibited "a capacity for change and for order hardly matched until the 19th century elsewhere in Europe," France in the 18th century retaining more provincial distinctions in law than did 11th-century England. The impressive central capability of the late Anglo-Saxon state was manifest in the extent of the fortress system, the regularity of large-scale surveys of land for fiscal ends, the sophisticated system of management of the coinage, the use of sealed writs for administrative purposes, and above all the capacity to tax. King Canute raised a larger income in real terms than did any of his medieval successors and, from his reign onward, disputes over landownership were frequently settled by reference to tax lists. Two Anglo-Saxon institutions left an enduring legacy for subsequent English state forms: the sheriff, already a royal official, and the shire court, within which large numbers of free men participated in the transaction of public business (Maitland [1908] 1963; Stubbs 1897). Some historians have gone still further, detecting in the late Anglo-Saxon realm an allegiance to what we would recognize as a nation (Wormald 1991).

The Norman conquest of 1066 allowed further consolidation of royal power. English feudalism was singular in that all alods (independently held estates) disappeared, and all land was nominally held of the king.

The crown took 20% of the land, the church another 25%, and a further 25% was dispersed among a dozen leading tenants-in-chief, but—critically—this was done in a way that minimized the possibility of large, territorially unified concentrations of seigneurial power that might challenge the crown. This too had Anglo-Saxon precedents (Wormald 1991). In most counties the king remained the single largest lay landowner. Monarchical strength in turn facilitated administrative innovation, though how much was consequential upon and how much preceded the Norman Conquest is now much debated (Campbell 1986). From the early 12th century, central branches of governance began to go "out of court," evolving into permanent institutions (though far from full-fledged Weberian bureaucracies), separate from the person of the king: first the Exchequer, then the major courts of Common Law (Common Pleas and King's Bench), then Chancery. Perry Anderson has quite rightly remarked that within medieval Europe, "justice was the ordinary name for power" (1974a, p. 153). In his terms, it was through the "superstructures" of jurisdiction that control of labor and its fruits was established. All the more significant, then, that a *Common* Law, administered in royal courts, was in place by the reign of Edward I, and that by the end of the 13th century, the assize system of justices of the central courts riding circuit throughout the realm twice a year, trying criminal cases and hearing civil suits at *nisi prius*, had been established.[4]

Entailed in all this is a separation of royal service, whether paid or not, from the fief. That in 1170 Henry II was able to fire all of his sheriffs tells much—"the king was master of those who governed in his name" (Bloch 1967, p. 430). Class relationships may have continued to rest on lords' command of what Anthony Giddens calls authoritative resources, but increasingly lords' authority became articulated through, and was correspondingly restrained by, a burgeoning machinery of state. The machinery undoubtedly served their interests as a class—Bloch argues that by this means English lords gained far more control over their peasants than did French (1967, p. 272)—but it restricted their powers as individual seigneurs. Influence and office were slowly becoming a more important source of power than one's following. By the close of the 16th century, the English aristocracy had largely ceased to be an actively military class. Conversely, and also from an early date, the crown itself was constrained by the routinized forms in which its authority became

[4] *Nisi prius* was in many ways a characteristically English institution in its blend of centralization and local involvement, a pattern I shall say more about below. Under this arrangement local juries were summoned to appear at the central courts on a certain day unless first (nisi prius) justices have appeared in their county to receive their verdicts and transmit them back to Westminster.

organized and through which its power was extended. Impeachment of royal ministers, to take but one example, was developed in the 1370s and 1380s.

Perhaps the most revealing illustration of the "precocity" of medieval English state formation, however, is its capacity for documentation. A critical dimension of supposedly modern power, stressed by social theorists from Weber to Foucault and Giddens, this welds individuals to the state as subjects, in both senses of that term. The *Domesday Book* survey of 1086, which itself probably had lesser Anglo-Saxon antecedents, is legendary. It took a mere two years to compile the book, rather less time than many a modern census. France was not to see a similar survey until 1327–28, and then six great fiefs were left out. Other records followed: Pipe Rolls of the Exchequer from 1131, Common Pleas from 1194, King's Bench from 1200, Parliamentary Rolls by 1340. Manuals of procedure, indubitable evidence of emergent routine, also appeared early. *Glanvill,* which describes the procedures of the royal courts, and the *Dialogue of the Exchequer* both date from the 1180s. The forms of writ that inaugurated the proceedings within the royal courts were established by the reign of Henry II. For Pollock and Maitland these represented "a new, authoritative and written formalism," which is "an English peculiarity, for the like of which we shall look in vain elsewhere" ([1898] 1968, p. 558). Year books that document pleas, evidence, and cases appear around the 1240s. A century later, a compilation is being made of the statutes of Parliament. From the 1530s on, all births, marriages, and deaths were compulsorily recorded in parish registers in order, as Thomas Cromwell put it, to determine "whether any person is our subject or no." This longevity of English records is surpassed in Europe only by the papacy.

This is clear evidence of formation of state; that is, separation of power from the person. Michael Clanchy analyzes writs, which bore the king's seal, using Weber's concept of routinization of charisma (1983, p. 156). Through these routines what Philip Abrams (1988) called "the idea of the state" is being given substance. The social identities of individuals are refashioned in terms of their membership in the national community that is defined by such exercise of power. Is the early date at which, in England, individuals assumed surnames (Laslett 1971, p. 47) one small sign of this, testimony to the requirements of identification? I do not know. But the body politic is no longer synonymous with the body of the king. It becomes possible to begin to conceive of the state as separable from the royal person, a prerequisite to holding the monarchy accountable to the community of citizens that it is claimed to represent. The rise of royal courts creates new rights that displace feudal obligations, private rights of the individual as a national citizen, which Marx and Weber both

see as basic to bourgeois civility (Sayer 1991). Surveys like *Domesday* and its many successors set up the state as the supreme arbitrator of entitlements to real property, giving a stability to possession that transcends legitimations of conquest or custom. Clanchy illustrates this beautifully in an anecdote of the *Quo Warranto* inquest of 1279. In Earl Warenne's confronting the commissioners with his rusty sword, protesting "This, my lords, is my warrant!" (Clanchy 1979, pp. 21–22), we have a poignant clash of legitimacies—several centuries before sociologists are apt to look for such conflicts of "tradition" and "modernity."

VI

Equally striking, and equally germane to the question of modes of legitimation, are the continuities in these institutions. This is as often a matter of form and symbol as it is of substance. Parliament, to take the most obvious example, was clearly not the same institution in 1450, 1650, and 1850. But form and symbol are themselves key elements of governance, "dignified" parts of the constitution that make for practical efficiency (Bagehot [1867] 1965). State formation is, among many other things, a cultural revolution, hinging upon the reconstruction of social identities (Corrigan and Sayer 1985). Assertion of continuity, or "invention of tradition" (Hobsbawm and Ranger 1983), is basic to this, situating the stately space in the flow of a national time linking the dead to the living and the unborn. In England, the fictions have been more than usually plausible.

Uniquely among the major states of Europe, what constitutes England as a geopolitical community has hardly altered since the country was first unified under the Wessex kings during the 9th and 10th centuries. Internally, the same is true. The county boundaries were little changed in England from before the Norman conquest until 1974, the best part of a thousand years. Parishes were likewise largely unaltered between the 12th and the 19th centuries. These are not just lines on maps, but locations of administration and power, a power that gave them a vitality as real communities. It was around the parish, at the lowest level, and the county at the higher, that governance was organized. The Elizabethan Poor Law and the 1662 Act of Settlement—two key instrumentalities for the construction of a labor market through regulation of vagrancy, remarked upon by both Marx (1967, chap. 28) and Weber (1966, pp. 227–28)—were administered through the parish, which, beginning with the Reformation, effectively became the lowest tier of government. The judicial system of Quarter Sessions and Assizes, the militia, and Parliamentary representation worked on the basis of counties, whose character as communities has been elucidated in the writings of Morrill (1974,

1976), Everitt (1966), and others. Such longevity is typical of many of England's legal and political institutions. The two original Courts of Common Law remained in existence up to 1875, when they became divisions of the High Court. The Assizes, regularized by statutes of 1299 and 1328, survived until 1971 and often sat in the place and on the day of the old Anglo-Saxon shire courts. Much of the substance of both the property and the criminal law laid down in the 12th century was also to endure to the 19th, though—characteristically—adjustments to changing times were made through legal fictions, while commentaries like Coke's and Blackstone's subtly altered what they purported merely to compile.

Of Parliament, more perhaps needs to be said. Its origins are murky and debated. By the reign of Edward I its convocation had become habitual. But, by around 1340, the twin bases of its power—legislative supremacy, and the right to grant taxes—were firmly being asserted, if not finally secured until (very much) later. It is a Whig fallacy to see Parliament as always having been opposed to the crown. The peculiarity of the English Parliament is rather that, instead of its being an institution that merely petitioned the crown, it developed as an integral element of the machinery of state, and this may have helped its later survival when, on the Continent, comparable bodies were falling into decay because of the consolidation of royal power (Elton 1974). There is a typically English pattern here, wherein the mutual constraints that parliamentary routines placed on the crown and notables at the same time empowered the former in its governance and the latter in prosecuting their private business. Parliament had three key distinctive features that enabled this. First, it was, from the beginning, a national assembly that had no regional competitors. Second, it was not organized by social estates. The clergy had no separate house after 1340, and (uniquely in Europe) the lesser nobility and gentry sat with the burgesses in the Commons. The House of Lords was in essence an enlarged King's Council, and the Commons did not represent a social class, commoners, but communes, the administrative cells of the kingdom. In this sense the English polity was not a conventional *Standesstaat*. Third, something made explicit in election writs from 1294 to 1872, M.P.'s were not delegates but representatives with *plena potestas* to commit their communities to laws and taxes: a power that is characteristic of a state (Cam 1970; Edwards 1970).

By the end of Elizabeth's reign, the English Parliament was the largest such body in Europe. It was also elected on a wider franchise than many. Long before that (I quote Chief Justice Thorpe, writing in 1366) it was held that Parliament represents "the body of the realm" and that its enactments were consequently binding upon all. The issue here is not whether or to what extent representation was democratic. In any modern sense it manifestly was not. The point, again, is that through routines of

state, a national community is being defined, membership in which socially identifies individuals as subjects. We need to take representation here in its other, Durkheimian sense: that of a *représentation collective* (Durkheim 1957, pp. 49–50). Certainly this political subjectivity is differentiated, by gender as well as by class. Most people remain "virtually represented." But it is in the time and space of the nation, nevertheless, that their social identities are increasingly delineated. The imagined community (Anderson 1983) thus constructed forcibly displaces other loci of identity and foci of loyalty, becoming the social terrain—the society, no less, in something like the modern sense of the word (Wolf 1988; Sayer 1991)—within which later political struggles are fought out.

Treason Acts, strengthened from the 1530s, broadcast the same message: one's first social identity cannot be "the Earl of Derby's man." As Maitland noted, this presumption has its roots in the peculiarities of medieval English feudalism, within which "it never becomes law that there is no political bond between men save the bond of tenure . . . whenever homage or fealty was done to any mesne lord, the tenant expressly saved the faith that he owed to his lord the King" (Maitland 1963, p. 161). As we move on into the Tudor period, which I lack the space to discuss more fully here (see Corrigan and Sayer 1985), these medieval fundaments are further consolidated. Critical is the nationalization of the church, precisely an act of state. Therewith accrues not only enormous wealth and a handy apparatus of patronage but also an extensive machinery of moral regulation and a vital repertoire of symbolic legitimation. In that same crucial decade, the 1530s, Parliament "revoked" to the crown most remaining jurisdictional liberties. Elizabeth substituted the royal arms for saintly images in parish churches, and her Privy Council set up Foxe's *Acts and Monuments,* with its tale of God's Englishmen (since Wycliff) battling the Antichrist, in every parish church in the land. The Anglican *Book of Common Prayer* nicely illustrates the Erastian intent in its commencing with an Act of Parliament, the Act of Uniformity. And in a whole series of areas, from the licensing of schoolmasters to the defining of heads of household, the density of government grows. There was, too, an upsurge in cartography under Elizabeth, with the imprimatur of Council. The state's making maps of the space it rules and represents is not just metaphorical.

VII

What might this state formation, which was undeniably impressive by the standards of the time, have to do with capitalism?

Most fundamental, I think, is the novel emergence of "the individual" as an autonomous subject, "freed" from the bonds of family, clan, or

estate and endowed with private rights and public representation. It is exactly such individuals who are agents of capitalist economy: independent subjects empowered to enter into contracts, including contracts for the purchase and sale of labor power. This is a major (if neglected) theme of classical social theories, about which I have written at length elsewhere (Sayer 1991). For Émile Durkheim, "it is only within the state that individualism is possible" (1957, p. 64). Karl Marx draws out the implication for capitalism more bluntly. "The attributes of the juridical person," he maintains, are "precisely [those] of the individual engaged in exchange" (1973, p. 246). We are not obliged to accept Marx's own view that it is exchange that creates this juridical person, rather than the other way about. In the case of England, I would argue, it was the exigencies of state that were primary. The individual was empowered by the forms in which royal authority was extended. Milsom puts it very well, speaking of the unintended long-term consequences of medieval Common Law. "The relationship between lords and tenants was subjected to the superior jurisdiction of royal courts and so became one of reciprocal private rights rather than of dependent allocation. . . . The transfer of jurisdiction made abstract rights out of claims to be allotted what was at management's disposal" (Milsom 1981, p. 5). The same can be argued of the parliamentary franchise vis-à-vis political feudalism, notwithstanding the rotten boroughs and bought elections that continued well into the 19th century. It created, if by no means in the ideal-typical shape of urban burghers, exactly Weber's "national citizen class" (1966, p. 249), and it did so from an early date. That such subjectivity did not yet embrace all (it was to remain propertied and gendered well into Weber's "age of iron") in no way denies the historical importance of its emergence.

As Weber emphasized, a uniform corpus of contractual and property law is essential to the existence of the modern permanent business enterprise. England enjoyed this from at least the 13th century, even if the forms of law were little codified. Weber suggests that in England a series of peculiarities of the Common Law fitted it well to capitalist needs in spite of its undoubted "irrationalities." These included the recruitment of judges and lawyers from the propertied and a costliness of procedures that effectively denied access to the courts to the poor (Weber 1978, pp. 891–92). This latter may be exaggerated. Relatively humble men did litigate, as well as serve on juries (and Weber's mention of the concentration of law courts in London forgets the provision for civil cases being tried in local assizes at *nisi prius*). This matters insofar as these legal routines defined subjectivities. That a yeoman could have recourse to the king's courts, and even participate in the administration of justice in the offices of constable, juryman, or churchwarden, served to enlarge and give substance to England as a political nation. The criminal law, on the

other hand, *was* heavily weighted against the accused. Not until 1692, for treasons, and 1836, for felonies, were defendants allowed the aid of counsel in preparing a case or cross-examining prosecution witnesses. This law was not only a medium through which rights were established for some, it was also an exemplary instrument for the disciplining of others—not least, as Marx traced (1967, pt. 8), with regard to vagrancy, trade unions, offences against property, and various forms of immorality alien to bourgeois proprieties.

A national law requires enforcement, which rests upon the state's monopolization of the legitimate use of violence. In England, although the aristocracy bore arms, by the Tudor period its social power, including its military power, was largely organized through state institutions. Elizabeth's 1573 reorganization of the militia confirmed the shift from following to office and influence as the basis of power. Maitland claimed that, at a much earlier date, English monarchy possessed "a counterpoise to the military system of feudalism" in the militia officered by the sheriffs, which dated from Anglo-Saxon times (1963, p. 162). That, in the Civil Wars, it was a Parliamentary army that challenged the king (and in the name of the law), underwrites this point. Along with this national uniformity of jurisdiction (and the power to levy taxes) go other forms of standardization that are conducive to capitalist enterprise: standardized currency, weights and measures, and customs' duties. In Germany, the *Zollverein* (customs union) of 1834 was a prelude to political unification. By 1700 England possessed the largest national market in Europe, to which Scotland was added by the Union of 1707. Hobsbawm—following Marx (1967, pp. 747–48) and Weber (1966, p. 230)—regards this home market as the tinder which the overseas trade, "backed by the systematic and aggressive help of government," ignited, and discovers in this troika the key to England's industrial revolution (Hobsbawm 1969, p. 50).

Fernand Braudel has written luminously on this. For him the "coherent, unified economic space" that forms a modern national economy is first of all "a *political* space, transformed by the state." "Only England," he argues, "managed this exploit at an early date." He emphasizes the very early removal of internal customs and tolls, the role of London as a go-between for the trade of the whole country, and the abundance of effective links of communication and transport. One important area he does not discuss is the grain market, which, from the Tudor period, was nationally organized by direct state regulation. All of these are circumstances contingent upon England's early state formation. Braudel goes on to say that within 18th-century France, by contrast, economic and political spaces did not coincide. There was a "maritime France" and a "continental France," and the latter "consistently controlled political power." Holland, in many other ways England's most

obvious rival for the title of "the classic ground," "occupied only a miniscule bit of territory, which was unable to feed its own population." For different reasons, but reasons which in both cases were as much political as narrowly economic, neither state constituted a sufficient national market for a capitalist take-off (Braudel 1977, pp. 99–111). Braudel also emphasizes England's political insularity (as I did above), which helped it "to protect its national market and burgeoning industries more successfully than any other European country" beginning with the 16th century (1977, p. 102). In fact, protectionism goes much further back: the first Navigation Act was passed in the reign of Richard II. As Weber says, "England was distinctively the original home of mercantilism" (1966, p. 256).

In these respects it is simply the organization of a society *as* a state that is propitious for capitalism. Equally important is what states can do, though here we do need to inquire into the class character and political agendas of those who staff or lobby them. State activities—as, again, both Marx (1967, pt. 8) and Weber (1966, pp. 227–28; 1978, p. 65; 1978, p. 930) detail—were fundamental to the construction of the social relationships upon which capitalism rests. In England property was privatized, and dispossessed peasants driven "onto the narrow path of the labor market" (Marx 1973, p. 507), through the use of state power. This is ground well-trodden by generations of social historians from Mantoux ([1928] 1964) to Snell (1985). To such commentaries I would simply add that the state's own interest in clear and documented title, if only for purposes of taxation, is likely a factor pushing toward more "modern" conceptions of what property is, quite independently of questions of class. Similarly, it was probably considerations of sound governance that impelled the draconian vagrancy legislation of later medieval and Tudor England of which Marx makes so much. That does not at all negate his observations as to its economic and moral consequences. Beyond this is the multifaceted project of civilizing society, or, as it might be seen from another viewpoint, organizing the subject (Foucault 1982; Sayer 1991), for which the state provided a battery of technologies of moral reformation that ranged from licensing of alehouses to provision of specimen sermons for "priests unlearn'd" (Corrigan and Sayer 1985).

The external facet of state is no less important. Weber (1966, p. 249) anticipates Wallerstein (1974) and Braudel (1977) in maintaining that a plurality of competing nation-states ensures that economics is not entirely subordinated to politics. Only within a world empire, like that of the Romans, the Chinese, or of Islam, can governments afford to buy social peace at the expense of economic enterprise. England was, for a long time, a national state vulnerable to world empires; Elizabeth's treatment of the privateering piracy of Drake and Hawkins tells much. Werner

Sombart long ago pointed out the important stimulus, both economic and technological, given to capitalism by state military competition. Related to this is the development of techniques of state-backed credit. England was to the fore in establishing both a national bank and a public debt, while during the 17th and 18th centuries, the use of warfare for commercial objectives was a norm of its state policy. Although neither Weber nor Marx locate the origin of capitalism in the looting of the rest of the world, both accept its contribution to capital formation where circumstances were otherwise favorable (e.g., Marx 1967, chap. 31; Marx [1864–65] 1971a, chap. 20; Weber 1966, chap. 26). Such looting could not have taken place to the extent that it did without gunboat diplomacy.

VIII

The way in which state formation may bear on three other English "peculiarities" deserves special mention. Many historians have remarked on the openness, in comparison with France or Germany, of the English elite from the Middle Ages on. Socially, and above all legally, the upper classes of England were both less differentiated internally than the French or German and were more penetrable by new men of property, including (consider Thomas Cromwell, Earl of Essex, son of a Putney blacksmith) property accrued through state service.[5] This view is not wholly refuted by the Stones's demonstration that fewer parvenu industrialists bought country estates in the 18th and 19th centuries than had once been believed (Stone and Stone 1984). The more important question has to do with the temper of the landed classes themselves; from very much earlier, they were ready not only to admit inferiors to their circle, but to engage in commercial activities themselves (or enter into convenient marriages) that might have compromised their honor elsewhere. For Max Weber, of course, the ethos of closed estates (*Stände*) is fundamentally opposed to that of capitalism. In England such an ethos was, at the very least, less entrenched than it is commonly thought to have been in much of Europe. By the 17th century, a landed class, of peers as well as gentry, whom Perry Anderson (1974b, p. 127) has rightly described as "unusually civilian in background, commercial in occupation and commoner in rank" stood in the very vanguard of the agricultural revolution that was to lay the foundation for 18th-century commercial and industrial capitalism.

---

[5] Perhaps the most powerful man in Henry VIII's England after the king himself, and according to Elton (1953, 1973) the real architect of the 'revolution in government' that accompanied the English reformation. His predecessor during the earlier part of Henry's reign, Cardinal Wolsey, also came from relatively humble beginnings.

This openness is hardly unconnected with the peculiarities of the English state. The royal courts were open to all free men and the Parliamentary franchise to all 40-shilling freeholders, independently of rank. Burghers sat with landed gentlemen in the Commons. Medieval kings succeeded in keeping open knighthood (it was a source of crown revenue), the early Tudors systematically buttressed the gentry against the magnates, while primogeniture and the eventual demilitarization of the aristocracy forced noble offspring into meaner pursuits. The younger son of a baron was a plain mister. Even the apparent exception, the peerage, confirms the argument. Never a true estate in the Continental sense, the peerage was circularly defined by summons to Parliament by royal writ; and as Maitland said, "It is well to observe how few were the privileges of peerage: how little of a caste was our estate of lords temporal" (Maitland 1963, p. 171). Nor did England come to possess any real equivalent of France's *noblesse de la robe*. In this adminstration of notables, there were simply fewer offices to farm.

Basic to England's agricultural revolution was the highly productive triad of landlord, tenant farmer, and landless laborer. Robert Brenner, in a very influential (if contentious) article (1985; cf. Aston and Philpin 1985), identifies this as *the* critical difference between England and France after the feudal crisis of the mid-14th century. In both cases, he says, the Black Death (which ravaged a society that was already undergoing a savage down phase of the Malthusian cycle) was met by a feudal reaction, taking the form, in England, of the 1349 and 1351 Statutes of Labourers. Nowhere within Western Europe was this reaction ultimately successful. But the English aristocracy was able to prevent its peasants' achieving security of tenure, despite an abundance of land and shortage of labor. In the long run this was to enable their eviction from the land, which allowed for a capitalist agriculture. In France, the crisis had a different outcome. The crown improved its position against the nobles by encouraging peasant smallholdings, a tactic that simultaneously limited seigneurial power and provided the monarchy with an independent tax base. But the price of this was high: "a long-term failure of agricultural productivity, [and] a corresponding inability to develop the home market" (Brenner 1976, p. 62). One can endorse this judgment without accepting Brenner's (much more debatable) general assumption that scale of investment per se is a critical determinant of agricultural productivity. As Alan Macfarlane points out (1987, pp. 176–77), however, Brenner never finally tells us *why* England and France differed in this regard. Part of the answer is likely to lie in the developments I have recounted above. As a class the English aristocracy were very much better placed to deal with peasant resistance than the French because their social power was organized in a national, statist form, within which institutions of

peasant communality—strong in certain ways as they might very well have been—were incorporated (one example is the organization of manorial courts); and this same state formation made any unholy alliance of crown and peasantry neither possible nor necessary.

The third factor worth highlighting here is the Northwestern European marriage system. This included (a) a late age (23 +) at first marriage for women, (b) a substantial proportion (10% +) of adults of both sexes remaining single, (c) the establishment of a new, and nuclear, household on marriage, and (d) a small age gap (3–5 years) between partners. So far as is now known, this pattern is distinctive to Northwest Europe, that is, to exactly the place where capitalism took root. And although historians disagree on its antiquity, which remains very much a matter of conjecture, there is no doubt that it goes back at least as far as the late 14th to 15th centuries; that is, to well *before* the rise of modern capitalism, as this is normally dated. Though not unique to England, it was there that this regime found its purest expression. Many historians (e.g., Laslett 1988; Macfarlane 1986, 1987; Levine 1989; Seccombe 1990) have contended that these connections are unlikely to be coincidental.

Under this marital regime fertility was less than half of what it would have been if all women 15–50 years old had been married. This allowed early modern England to escape from the Malthusian cycles found in much of Europe, while retaining the potential for a flexible response to secular booms, as in the late 18th century, providing a supply of new labor in times of sustained economic expansion. It left a margin for savings that was not continually being eroded by population pressure and provided incentives to save—even for those well down the social scale—by making marriage contingent upon the economic resources to set up a new household. It also made a contribution to capital formation by decreasing the proportion of youthful dependents and widows and enhancing the productivity of female labor. It made available a pool of youthful wage labor of both sexes and redistributed this labor to the more productive holdings. Such service-in-husbandry was well established in medieval England. As important as these directly economic effects of this marital regime, however, are the cultural corollaries emphasized by Macfarlane (1986). This system, he argues, severs the link found in many "traditional" societies between adulthood and parenthood, fostering individualism as a cultural norm and calculation as an orientation toward action—eminently Weberian attitudes.

Macfarlane is also clear, however, that the marriage system could only have this result, given certain other conditions. These are those theorized by Thomas Malthus himself: an acquisitive ethos, the opportunity of social mobility, a generally elevated standard of living, and security in the fruits of one's labor. It makes sense not to invest in children only

where, on the one hand, there is no compelling political motive (whether security, protection, or influence based on kinship) for procreation and, on the other, where alternative investments are an economically superior risk. For Macfarlane, England uniquely met these conditions for a variety of reasons. Some of these, like Christian attitudes toward both sexuality and marriage, were part of a wider European heritage. But distinctive to England was the "powerful, unified, political system . . . built up by the Anglo-Saxon kings and consolidated by the Normans and Angevins," in which "public peace and the control of violence were in the hands of chosen officials rather than the family's" (Macfarlane 1986, p. 335). Here, too, then, in what may prove to be a vital missing link in explaining "the European miracle," England's early state formation seems to have been a sine qua non.

IX

In the light of these arguments, it is surprising that England's state formation should have received so little attention in the literature on the rise of capitalism, particularly when Skocpol (1979), Giddens (1985), Mann (1986), and others are pleading eloquently for the need to "bring the state back in." I suspect the reason for this lies in the tyranny of ideal-types. It is not easy for those trained to distinguish between tradition and modernity to view Shakespearian parish constables, a hereditary upper chamber, or the arcane rituals surrounding the British crown as sinews of a bourgeois state. I believe this to be fundamentally mistaken. England's polity was to turn out to be remarkably well suited to the demands of capitalism. And, I would argue, it is precisely its archaic, unrationalized, traditionalist qualities—its "logical incongruities"—that go a long way toward explaining why. This brings me to the final "peculiarity" on which I wish to focus here, the "amateurism" of England's political and legal institutions. Quite as significant as the fact of national state formation were the *forms* this state formation took, and who they empowered to do what

Let me exemplify one such governing institution, that of the Justices of the Peace (J.P.'s). It is an institution on which Weber (1978, pp. 1059–64) makes some telling observations, entirely congruent with my own. In many ways the J.P.'s "symbolize the polity of England" (Gleason 1969, p. 1). Originating in the 13th century as knights commissioned to keep the peace, they came, increasingly, to acquire juridical functions. These were placed on a statutory basis under Edward III. Justices of the Peace had full jurisdiction, in Quarter Sessions, over all misdemeanors and felonies (though from the Tudor period they were directed to leave more difficult cases for the assize judges) and could try more minor crimes

summarily. Over the centuries, and especially under the Tudors, they added to this a very wide range of administrative responsibilities, from upkeep of highways to licensing of corn merchants. For half a millenium they regulated wages, county by county and trade by trade. They scrutinized parish accounts and oversaw Overseers of the Poor. Until the 19th century, for most English people and most practical purposes, J.P.'s *were* the state.

What is most interesting, and most characteristic, about this institution is the nice balance of center and locality, amateurs and professionals, which it epitomizes. It *was* self-government, *but* at the king's command. Justices of the Peace were drawn from the propertied gentlemen of their county. From 1439 they were required to have land tenements to the value of £20 per annum, 10 times the property qualification for the franchise, and this was raised to £100 in 1732. The day-to-day governance of the realm was thus largely devolved to local elites acting in an unpaid, amateur capacity, and this juridical role, of course, also empowered them in their private stations as landlords. Justices had considerable latitude in the way they chose to administer their counties. At the same time, they were scrutinized (with varying intensity according to the period) from above. The law which they enforced was not a seigneurial *haute justice,* but Common Law, and the Clerks of the Peace (who were legally trained) advised them and recorded their decisions. Justices of the Peace were personally commissioned by the crown, and plentiful petitions to Parliament requesting that they be elected were refused (Maitland 1963, p. 207). The bench was from time to time purged, as it was of recusants, under Elizabeth, or Tories, under Walpole. From the 14th century onward, J.P.'s operated within a strict framework of statutory regulation, in which they were personally liable for maladministration. Council closely watched their activities, as did the Justices of Assize; under Elizabeth, for instance, they imposed the oath of supremacy on all J.P.'s.

A similar pattern can be observed of other distinctively English institutions. A peculiarity of the Common Law was the jury system, which involved local men originally chosen not for their ignorance of the facts and individuals on trial but for their knowledge of them. Deputy Lieutenants, who commanded the militia, were appointed by Lord Lieutenants of the counties, but the crown had a veto on their choice. Parliament itself brought those whom Hexter (1961) calls "country magnates" to Westminster to vote taxes and assent to legislation, giving them a voice and role in national governance, and, as Weber noted, its provision of borough seats ensured early representation of urban interests in (rather than against) the community of state. But Parliament has throughout its history been subject to executive manipulation, and taxes were not for the most part farmed out.

This structure evidently empowers local notables, both publicly and privately. But it does so within and through the arena of state that these institutions define. Both points are important. Though local elites have broad legal and governmental powers, this is by no means a parcelization of sovereignty. Though the center surveys and supervises, this is very far from an absolutism. This polity is a formation sui generis and implicit in it is a double constraint. From the Middle Ages on, even the mightiest of subjects are increasingly impelled, by what they stand to gain as much as by what they might lose, to channel their ambitions and conflicts through the medium of state forms. But as those forms attain definition, stability, and regularity, the crown too is restrained by them, as the Stuart kings were to discover. In the famous contrast which Sir John Fortescue drew in his *Learned commendation of the politique laws of England,* a work written in the 1460s, we have, not a *dominium regale* but a *dominium politicum et regale*—or, in more modern terms, something approaching a constitutional monarchy.

X

Such locally devolved, yet centrally regulated, institutions of governance were arguably very much better suited to the needs of a developing capitalism than either a more bureaucratized, or a more bourgeois-democratic, polity might have been. They were, to begin with, cheap, since there was little burdensome staff of permanent officials to finance, whether by salaries or levies on subjects imposed by officeholders themselves. This bears on the ability, which North and Thomas (1973) maintain is critical to capitalist development, to maintain an optimal balance between private and social returns, avoiding overtaxation and sacrifices of long-term investment to current revenue. Equally important, these state forms left men of property largely free to "improve" their possessions as they thought fit, with minimal interference from above.

To be sure, various "predatory complexes" (Thompson 1978) would from time to time occupy the commanding heights of state, and Court/Country tensions have been a recurrent feature of English history from the Household versus *Curia* ambivalences of the Middle Ages onward. But because of the very practical dependence of central governance on the involvement of local notables, the center ultimately had little option but to accommodate to their demands, and English state forms were well suited to accomplish this. Quarter Sessions, Assizes, and Parliament were regularized occasions for the mutual adjustment of central policies to local circumstances, on the basis of an effective two-way information flow. Such institutions drew a wide circle of notables into the broad community of state—even, in their humble way, down to the village

elites of churchwardens, constables, and Quarter Sessions' jurymen, who by the 17th century were often substantial commercial farmers (Wrightson and Levine 1979, Wrightson 1982)—giving them channels of national political expression. This national location in turn authorized their personal standing in county and parish, cloaking private interests in all the majesty of state. Central to these interests, from the Middle Ages on, were the prerogatives of property, even as the economic character of the latter shifted in capitalist directions.

This amateurism also entailed the inscription of localized hierarchy as an instrument of national administration. These state forms mobilized informal networks of kinship and clientage as means of governance, but—again we see the duality—made access to the means of governance itself a critical element in family fortunes. This too seems "archaic," not to say plainly corrupt: a species of corruption, in this case, that extended well into the 20th century (Corrigan and Sayer 1985). But it had its advantages. The personalized character of power that is basic to feudal and patrimonial rule, beautifully brought out for late medieval England by Colin Richmond (1988), remains, but becomes an animating and legitimating component of a machinery of state that is modern in its reach, coherence, and capabilities. Such institutions facilitated not only the cohesion of elites but also the moral surveillance and policing of the laboring poor.

Mesmerized by "free" labor—for both Marx and Weber the *differentia specifica* of modern capitalism—we can overlook the degree to which the imagery (and, until 1875, statutes) of master and servant long continued to structure terms of employment in England. This is most evident in the tied cottages, the annual hiring fairs, and partial payment-in-kind of farm laborers (known as servants-in-husbandry) or in the domestic service in which, as late as 1851, more people (overwhelmingly women) were employed, as Marx (1963, pp. 200–201) sourly noted, than in factories. Typically, he saw this in terms of the bourgeoisie improperly aping "feudal" mores, but such anachronisms were also informative of the industry on display at the Great Exhibition in the Crystal Palace that same year. Extended family groups followed coalmasters from one mine to another, and these masters "exercised a considerable degree of social control over their workforce" (Nair 1988, p. 252). Douglas Hay (1975) brings out the embeddedness of personal paternalism in tempering the "bloody code" of the 18th-century criminal law. A good character reference from his master could save a felon from the gallows. Even the supposedly rational organs of moral discipline of the late 19th century— the schools, charities, and libraries—remained in many instances the gifts of local notables, modifying, but by no means eliminating, the older patriarchal ethos.

Preindustrial (but long capitalist) England was not a "one-class" society, save perhaps in Peter Laslett's (1971) sense of a society in which only one social group, its rulers, was self-consciously organized as a class. But Keith Wrightson's (1986) description of it as constituted by both horizontal relations of class and vertical relations of paternalism, the latter both orchestrating and symbolizing the former, is very much to the point. These institutions of state (and Abrams's "idea of the state" [1988]) were shot through with this imagery. The state was nationally projected and locally administered as a patriarchal family writ large (Amussen 1988; Corrigan and Sayer 1985), within which the masters virtually represented their multiple personal dependents. To be sure, the forms of social discipline changed over time, notably with the growth of industry and towns. But this was, from the point of view of the rise of capitalism, late in the day, and even then (like the industrial revolution itself) did not go smoothly or happen overnight. The legal definition of murder of a man by his wife, or of a master by his servant, as petty treason (which was punishable by death by burning instead of hanging) was removed from the statute book only in 1828. That these were crimes against the state perfectly captures the mirroring of England's domestic and state patriarchies.

Finally, I would suggest, these state forms came to acquire substantial legitimacy, authorizing in the name of the nation an exercise of power that long remained, in both its agency and the interests it served, that of private gentlemen. This legitimacy is not unconnected with either the patriarchy or the antiquity of these institutions, nor with their claim to represent the nation they did much to construct. England is striking (although some other old polities, like the United States, share this quality) in the degree to which nationality is symbolized by the paraphernalia of state, which take on a homely, familiar aspect: the Royal Family, Big Ben, the friendly (and "unarmed") London bobby. The irrationality of English ways of doing things itself partakes of this. For Burke, the beauty of the English constitution is that it is not derived by rational deliberation, but distills all the experience and wisdom of the ages. Political theorists from Hobbes to Durkheim have dwelt on the way state power is akin to that of religion. These are totems that signify who we are—even if that identity is one we are constrained to assume—and they inspire, as well as awe, a certain reverence.

Perhaps the ultimate testimony to the resilience of English political forms lies in that way that they informed and embraced opposition. Aylmer's observation that the Levellers' aspirations "owed a great deal to the notion and to the partial practice of self-government, locally if not nationally, and to the tradition, even if it were partly a hollow one, of the individual's rights at law" (Aylmer 1975, p. 13) holds for many later

English social movements (see, e.g., Thompson 1968, chap. 1). Their political object was more often a full inclusion within the communality of state than its revolutionary overthrow. At stake here is something much more profound than ideological hegemony. As Marc Bloch once put it, "in an old society, cemented by a' shared culture, the humble are . . . always, for good or ill, constrained to make common cause with the mighty" (1968, p. 141). The issue is *not* an absence of social and political struggles, but the terrain on which they came to be waged; that landscape long mapped by the English state as a unitary society in which all, differentially, partake.

As the nature of English society changed, these institutions proved themselves flexible enough to adjust and, indeed, to serve as the vehicles for articulation of changing class interests over time. This, I believe, is a function of their ancient rootedness in the political nation, that nation which extended from parish to Parliament and knitted the one to the other, enshrining in its polity the patriarchy of locale. Chirot's observation apropos parliaments, that "what had seemed to be a 17th century remnant . . . proved to be more flexible and conducive to progress than the newer, more centralized and bureaucratized ways of Europe's absolutist monarchs" (Chirot 1985, p. 186) applies more generally. English institutions were permissive enough to allow for private enterprise to flourish within the civil society they had molded, and adaptable enough to accommodate the new social forces which that enterprise unleashed. Though not without conflict, this is what happened in the end. Manchester and Birmingham achieved Parliamentary representation consistent with their populations. Industrialists joined clergymen and gentry on the county bench. New instrumentalities of regulation, such as inspection (Corrigan and Sayer 1985, chap. 6), were developed, which still echoed the "traditional" pattern of center/locality relations. Old forms, like the Royal Commission, were adapted to modern purposes but still manned by the voluntary legions of "The Great and The Good" (Ashforth 1990). In the end working men, and still later, women, were enfranchised. The sky did not fall. But nor, at any point, was the essential capitalist freedom, that of private enterprise, ever seriously jeopardized.

XI

I have not argued here that state formation was the cause of capitalism, in England or anywhere else. That would be fatuous. Nor have I attended overmuch to the proximate causes of economic take-off. My concern has rather been with the multiple ways in which—to stand Karl Marx's "guiding thread" ([1859] 1971b) on its head—English state for-

mation, over the very *longue durée,* molded a civil society in which capitalist economy was possible. Critical in this was the early unification of England as a national state. But equally significant was the "peculiar" character of the political and legal institutions through which this national unification was accomplished and in whose continuities, real and imagined, society came to be represented. It is in this that England most evidently deviates from Marxist and Weberian ideal-types. My contention here has been that it was precisely these deviations that were conducive to the success of its pioneering capitalism.

These institutions of state directly, and from a very early date, empowered a wide spectrum of Englishmen of property against encroachments, whether from above or below. They proved sufficiently flexible, by virtue of their rootedness in locality and interest, to accommodate secular economic and social changes in the connotation of property. They long subjected the populace to a social discipline that was local, personal, and patriarchal, and the more effective for all that. And they legitimated all these things, superbly, in the name of an English nationality claiming to embrace everyone in a common and ancient polity and culture. From the point of view of its contribution to the rise of capitalism, there would appear to be every reason for regarding such a polity as paradigmatic, rather than peculiar.

I intend this in a strong sense. England undoubtedly was distinctive in many of the particulars I have narrated here, but its distinction was more often of degree than of kind. It only appears as peculiar, as wholly singular, exceptional, unique, in the light of our ideal-typical expectations of what is normal, that lead us one-sidedly to focus on other factors when looking at other polities. I suspect that were we to treat England as a paradigm, and ransack other histories for functional equivalents of what I have argued in this essay was most conducive in English state formation to the rise of capitalism, we would not come up empty-handed. Recent "revisionist" studies of the French Revolution (Furet 1981, Schama 1989) offer support for this view. But the most intriguing parallels may lie on the other side of the modern world. Taiwan and Korea are hardly paragons of rational-legal administration or bourgeois democracy. And the most spectacular capitalist success story of the 20th century is an island state, characterized by long-standing political integrity and a paternal social ethos, in which what became the Mitsubishi Corporation was established through the state, following an imperial Restoration, by the "feudal lords" of Satsuma. This has not been an essay in comparative sociology, and comparison would be needed to test its theses. But it bears, I think, on more than the peculiarities of the English. What is in question is rather the standpoints from which, incongruously, "the classic ground" is deemed peculiar in the first place.

REFERENCES

Abrams, P. 1988. "Notes on the Difficulty of Studying the State." *Journal of Historical Sociology* 1(1):58–89.

Amussen, S. D. 1988. *An Ordered Society: Gender and Class in Early Modern England*. New York: Basil Blackwell.

Anderson, B. 1983. *Imagined Communities*. London: Verso.

Anderson, P. 1974a. *Passages from Antiquity to Feudalism*. London: New Left.

———. 1974b. *Lineages of the Absolutist State*. London: New Left.

Ashforth, A. 1990. "Reckoning Schemes of Legitimation: On Commissions of Inquiry as Power/Knowledge Forms." *Journal of Historical Sociology* 3(1):1–22.

Aston, T., and C. Philpin, eds. 1985. *The Brenner Debate*. Cambridge: Cambridge University Press.

Aylmer, G. E. 1974. *The King's Servants*, rev. ed. London: Routledge & Kegan Paul.

———. 1975. *The Levellers in the English Revolution*. London. Thames & Hudson.

———. 1986. *Rebellion or Revolution: England 1640–1660*. New York: Oxford University Press.

———. 1990. "The Peculiarities of the English State." *Journal of Historical Sociology* 3(2):91–108.

Bagehot, W. (1867) 1965. *The English Constitution*, edited by R. Crossman. London: Fontana.

Bloch, M. 1967. *Feudal Society*. London: Routledge & Kegan Paul.

———. 1968. *Strange Defeat*. New York: Norton.

Braudel, F. 1977. *Afterthoughts on Material Civilization and Capitalism*. Baltimore. Johns Hopkins University Press.

Brenner, R. 1985. "Agrarian Class Structure and Economic Development in Pre-Industrial Europe." Pp. 10–63 in *The Brenner Debate*, edited by T. Aston and C. Philpin. Cambridge: Cambridge University Press.

Brewer, J 1989. *Sinews of Power*. London. Unwin Hyman.

Cam, H. 1970. "The Theory and Practice of Representation in Medieval England." Pp. 263–89 in *Historical Studies of the English Parliament*, vol. 1. Edited by E. Fryde and E. Miller. Cambridge. Cambridge University Press.

Campbell, J. 1975. "Observations on English Government from the 10th to the 12th Centuries." *Transactions of the Royal Historical Society*, 5th ser., 25. 39–54.

———. 1986. *Essays in Anglo-Saxon History*. Roncevert, W.V.: Hambledon.

Chirot, D. 1985. "The Rise of the West." *American Sociological Review* 50:181–95.

Clanchy, M. 1979. *From Memory to Written Record*. London: Edward Arnold.

———. 1983. *England and Its Rulers, 1066–1272*. London: Fontana

Clark, J. C. D. 1986. *Revolution and Rebellion: State and Society in England in the 17th and 18th Centuries*. Cambridge: Cambridge University Press.

Corrigan, P., and D. Sayer. 1985. *The Great Arch: English State Formation as Cultural Revolution*. New York: Basil Blackwell.

Durkheim, É. 1957. *Professional Ethics and Civic Morals*. London: Routledge & Kegan Paul.

Edwards, J. G. 1970. "The Plena Potestas of English Parliamentary Representatives." Pp. 136–49 in *Historical Studies of the English Parliament*, vol. 1. Edited by E. Fryde and E. Miller. Cambridge: Cambridge University Press.

Elton, G. 1953. *The Tudor Revolution in Government*. Cambridge: Cambridge University Press.

———. 1974. *Studies in Tudor and Stuart Politics and Government*. Vol. 2, "Parliament, Political Thought." Cambridge: Cambridge University Press.

Engels, F. 1962. "On Certain Peculiarities of the Economic and Political Development

of England." P. 529 in *On Britain* by K. Marx and F. Engels. Moscow. Foreign Languages Publishing.

Everitt, A. 1966. *The Community of Kent and the Great Rebellion*. Leicester: Leicester University Press.

Furet, F 1981 *Interpreting the French Revolution*. Cambridge: Cambridge University Press.

Giddens, A 1985. *The Nation-State and Violence* Berkeley and Los Angeles: University of California Press

Gleason, J 1969 *Justices of the Peace in England, 1558–1640*. Oxford: Clarendon.

Goldstone, J. 1988. "East and West in the 17th Century Political Crises in Stuart England, Ottoman Turkey and Ming China." *Comparative Studies in Society and History* 88.103–42.

Hay, D. 1975. "Property, Authority and the Criminal Law " Pp. 17–64 in *Albion's Fatal Tree*, edited by D. Hay et al. London: Allen Lane.

Hexter, J. 1961 *Reappraisals in History* London. Longmans.

Hill, C. 1955. *The English Revolution 1640*, 3d ed. London. Lawrence & Wishart.

———. 1974. "Protestantism and the Rise of Capitalism " Pp. 81–102 in *Change and Continuity in 17th Century England* by C. Hill. London: Weidenfeld & Nicholson.

———. 1980. "A Bourgeois Revolution?" Pp 109–39 in *Three British Revolutions*, edited by J. Pocock. Princeton, N.J.: Princeton University Press

———. 1981 "Parliament and People in 17th Century England " *Past and Present* no. 92:100–24.

Hilton, R., ed. 1978. *The Transition from Feudalism to Capitalism*. London: Verso.

Hirst, D. 1981. "Revisionism Revised. The Place of Principle." *Past and Present* no. 92:79–90.

Hobsbawm, E. 1969. *Industry and Empire*. London. Penguin.

Hobsbawm, E., and T. Ranger, eds. 1983. *The Invention of Tradition*. Cambridge: Cambridge University Press.

Kennedy, P. 1988. *The Rise and Fall of the Great Powers*. London: Unwin Hyman

Laslett, P. 1971. *The World We Have Lost*, 2d ed. London: Methuen.

——— 1988. "The European Family and Early Industrialization." In *Europe and the Rise of Capitalism*, edited by J. Baechler, J. A. Hall, and M. Mann. New York: Basil Blackwell.

Levine, D. 1989 "Recombinant Family Formation Strategies." *Journal of Historical Sociology* 2(2):89–115.

Lowith, K. 1982. *Max Weber and Karl Marx*. Trans. by T. Bottomore and W. Outhwaite. London: Allen & Unwin.

Macfarlane, A. 1986. *Marriage and Love in England, 1300–1840*. New York: Basil Blackwell.

——— 1987 *The Culture of Capitalism*. New York Basil Blackwell.

Maitland, F W (1908) 1963 *The Constitutional History of England*. Cambridge. Cambridge University Press.

Mann, M. 1986. *The Sources of Social Power*. Vol. 1, "A History of Power from the Beginning to A D. 1760." New York. Cambridge University Press.

Manning, B. 1978. *The English People and the English Revolution*. London: Penguin.

Mantoux, P. (1928) 1964. *The Industrial Revolution in the 18th Century* London: Methuen

Marx, K (1861–63) 1963 *Theories of Surplus Value*, vol. 1 Moscow. Progress.

———. (1867) 1967. *Capital*, vol. 1. London: Lawrence & Wishart.

———. (1864–65) 1971a *Capital*, vol 3 Moscow. Progress

———. (1859) 1971b Preface to *A Contribution to the Critique of Political Economy* London: Lawrence & Wishart.

———. (1857–58) 1973. *Grundrisse*. London. Penguin.

Marx, K., and F. Engels. (1845–46) 1976a. *Collected Works*. Vol. 5, *The German Ideology*. London. Lawrence & Wishart.
———. (1847–48) 1976b. *Collected Works*. Vol. 6, *Manifesto of the Communist Party*. London: Lawrence & Wishart.
———. (1855) 1980. *Collected Works*. Vol. 14, *The British Constitution*. London: Lawrence & Wishart.
——— (1855) 1986. *Collected Works*. Vol. 22, *The Civil War In France*. First Draft. London: Lawrence & Wishart
Milsom, S. 1981. "The Nature of Blackstone's Achievement." London: Selden Society pamphlet.
Morrill, J. S. 1974. *Cheshire 1630–1660: County Government and Society during the "English Revolution."* Oxford. Oxford University Press.
——— 1976. *Revolt of the Provinces*. London. Longmans.
Nair, G 1988 *Highley: The Development of a Community, 1550–1880* New York: Basil Blackwell
North, D. C., and R. P. Thomas. 1973. *The Rise of the Western World* Cambridge: Cambridge University Press.
Pirenne, H. (1925) 1974. *Medieval Cities: Their Origins and the Revival of Trade*. Princeton, N.J.. Princeton University Press.
Pollock, F., and F. W. Maitland. (1898) 1968. *History of English Law*, vol. 2. 2d ed. Cambridge. Cambridge University Press.
Richmond, C. 1988 "Hand and Mouth. Information Gathering and Use in England in the Later Middle Ages." *Journal of Historical Sociology* 1(3):233–52.
Russell, C. 1979. *Parliaments and English Politics, 1621–1629*. Oxford· Oxford University Press
Sayer, D 1987. *The Violence of Abstraction*. New York. Blackwell.
Sayer, D., ed. 1989. *Readings from Karl Marx*. New York· Routledge.
———. 1991 *Capitalism and Modernity: An Excursus on Marx and Weber*. New York: Routledge.
Schama, S. 1989 *Citizens*. New York: Knopf.
Seccombe, W. 1990. "The Western European Marriage Pattern in Historical Perspective A Response to David Levine." *Journal of Historical Sociology* 3(1):50–74
Segal, D 1988. "Nationalism, Comparatively Speaking." *Journal of Historical Sociology* 1(3).301–21.
Skocpol, T. 1979. *States and Social Revolutions*. Cambridge: Cambridge University Press
Smiles, S. (1859) 1986. *Self-Help, with Illustrations of Conduct and Perseverance*, edited by Sir K Joseph, Bt, M.P. London. Penguin.
Snell, K. D. M. 1985. *Annals of the Labouring Poor: Social Change and Agrarian England, 1660–1900*. Cambridge: Cambridge University Press.
Steedman, C. 1984. *Policing the Victorian Community: The Formation of English Provincial Police Forces*. London: Routledge & Kegan Paul.
Stone, L., ed 1965. *Social Change and Revolution in England*. London Longmans.
Stone, L., and J. Stone. 1984 *An Open Elite? England, 1540–1880*. Oxford. Clarendon.
Stubbs, W 1897. *The Constitutional History of England*. 3 vols , 6th ed. Oxford. Clarendon.
Sweezy, P. 1978. "A Critique" [of Maurice Dobb] Pp. 33–56 in *The Transition from Feudalism to Capitalism*, edited by R. Hilton London: Verso
Tawney, R. 1938. *Religion and the Rise of Capitalism* London: Penguin.
———. 1941a "Harrington's Interpretation of his Age." *Proceedings of the British Academy* 27:199–224.
———. 1941b. "The Rise of the Gentry, 1558–1640." *Economic History Review* 11.1–38.

Thompson, E. P. 1968. *The Making of the English Working Class*. Rev. ed. London. Penguin.

———. 1978. "Peculiarities of the English." Pp. 35–91 in *The Poverty of Theory and Other Essays* by E. P. Thompson: London: Merlin.

Tilly, C. 1990. *Coercion, Capital and European States,* AD *990-1990*. Cambridge, Mass.: Basil Blackwell.

Wallerstein, I. 1974. *The Modern World System,* vol. 1. New York: Academic Press.

Weber, M. (1904–18) 1949. *Methodology of the Social Sciences*. Edited by E. A. Shils and H. A. Finch Glencoe, Ill.: Free Press.

———. (1920–23) 1966. *General Economic History* Edited from lecture scripts by S Hellman and M. Palyi. New York: Collier.

———. 1970. *From Max Weber*. Edited by H Gerth and C. Wright Mills. London. Routledge & Kegan Paul.

———. (1904–5). 1974. *The Protestant Ethic and the Spirit of Capitalism*. London: Allen & Unwin.

———. (1910–20) 1978. *Economy and Society*. Edited by G. Roth and C. Wittich, 2 vols. Berkeley and Los Angeles: University of California Press.

Wolf, E. 1988. "Inventing Society." *American Ethnologist* 15(4):752–61.

Wormald, P. 1991. *"Engla Lond:* The Making of an 'Alternative Allegiance.'" Paper presented at the 10th annual meeting of Discussion Group on the English State, St. Peter's College, Oxford, April 1991

Wrightson, K. 1982. *English Society, 1580–1680*. London: Hutchinson.

———. 1986. "The Social Order of Early Modern England " In *The World We Have Gained,* edited by L. Bonfield, R M Smith, and K. Wrightson. Oxford: Oxford University Press.

Wrightson, K., and D. Levine. 1979. *Poverty and Piety in an English Village: Terling, 1525–1700*. New York: Academic Press.

Zagorin, P. 1970. *The Court and the Country: The Beginning of the English Revolution of the mid-17th Century*. London. Routledge & Kegan Paul.

The Origins of Private Social Insurance: Public Policy and Fringe Benefits in America, 1920–1950[1]

Frank R. Dobbin
Princeton University

How did the American system of private, employment-related pension and health insurance arise? Data on corporate fringe-benefit programs during the second quarter of the 20th century contradict the received wisdom that benefits rose in response to wartime federal policy changes and industrial factors. Instead it appears that public policies such as the Wagner Act and Social Security led to union and business support for private insurance, which in turn spurred the growth of fringe benefits. The historical record suggests that neoinstitutional and conflict approaches must be synthesized to explain the expansion of fringe benefits: institutional factors influenced organizational outcomes by affecting interest group goals.

INTRODUCTION

Over the past century the United States has developed an elaborate system of private, employment-related insurance to protect Americans against income loss due to illness and old age. Today over half of all full-time employees in the private sector are covered by employer-provided health insurance and pension benefits (Rein 1982, p. 132). In 1980, the private sector paid for 57% of total U.S. health expenditures, an amount nearly 3.5 times greater than the average (17%) private-sector contribution to health care costs among Western European nations.[2]

[2] The data on combined pension and health coverage includes only full-year nonfarm employees. The countries for which comparative data were available are Austria, Belgium, Denmark, Finland, France, Germany, Ireland, Italy, the Netherlands, Norway, Sweden, Switzerland, and the United Kingdom (Esping-Andersen 1990, p. 69).

Likewise, the private sector provided 21% of total pension benefits in the United States, more than double the average private-sector contribution (9%) in Europe (Esping-Andersen 1990, p. 85). In an effort to understand the origins of America's system of employment-related income protections, this article examines the decline of industrial welfare work and the rise of corporate fringe-benefit programs in the second quarter of the 20th century.

Forms of Employment-related Coverage

Welfare work.—Pension and health coverage appeared along with other forms of "welfare work," such as company housing and recreational programs, in the teens and twenties in firms like General Electric, U.S. Steel, Pullman, and International Harvester. Employer-financed welfare programs were unsecured, and, because they were popularized as a means to increase worker loyalty, benefits were commonly denied to known union activists (Quadagno 1988, chap. 4). Employers paid for early "informal" pension programs from current income, and mutual benefit associations maintained contributory employee-financed funds that paid benefits to injured and ill employees and sometimes to their survivors (Brandes 1976). Such associations had been established as early as the 1860s in dangerous industries, and in the 1920s they became an integral part of company welfare work. Like informal pension schemes, most mutual benefit associations could be terminated by the employer at will. While some employers used welfare work to fend off unions, others simply saw it as a way to maintain the work force. As one Midwestern businessman put it, "When I keep a horse and I find him a clean stable and good food I am not doing anything philanthropic for my horse" (Brandes 1976, p. 31).

Health and pension insurance.—Between the twenties and the fifties, firms installed health and pension insurance plans that put coverage on a sound actuarial basis and guaranteed that benefits would be paid. Group pension insurance was financed by joint employer-employee contributions to a private insurance company program, and benefits were usually calculated on the basis of years of service. Group accident/sickness insurance guaranteed a flat daily payment for each day of work missed on account of illness or injury—typically two-thirds of normal wages (James 1947, p. 263; National Industrial Conference Board [NICB] 1934, p. 13). In many settings this early form of health coverage was later replaced or supplemented by medical and hospitalization coverage.

Perspectives on Employment-related Coverage

Industrial arguments.—A number of analysts have linked the rise of group insurance to industrial factors. First. some trace insurance schemes to efforts to quell turnover when competition for workers heated up in the late thirties and forties—health and pension benefits based on job tenure were expected to discourage job changes (see Slichter 1961). Second, labor segmentation theorists suggest that turnover was most costly in capital-intensive, core industries that relied on firm-specific skills, and that those industries increased wages and benefits to induce long-term employment (Doeringer and Piore 1971, Hodson 1978; Edwards 1979, p. 142). In support of this view Gordon, Edwards, and Reich (1982, pp. 195–212) find sectoral wage differences as early as 1914 and argue that today's labor-market segments and associated remuneration differences had crystallized by 1950. Third, others argue that union gains after the Wagner Act stimulated the growth of fringe benefits, either because unions sought new victories in contract negotiations or because employers voluntarily adopted fringe benefits to thwart organizing efforts (Bernstein 1972; Jacoby 1985). Finally, organizational theorists have argued that growth in firm size leads to the formalization of employment relations (Blau and Schoenherr 1971; Pugh et al. 1969), and analysts indeed date the rise of formal, insured types of health and pension coverage to the war years, when firms grew dramatically in size. Arguments broadly similar to these have been made about the rise and decline of welfare work. Analysts suggest that welfare work rose in response to union activism in the twenties and declined in the early thirties when the Depression undermined union power and dampened labor turnover (Brandes 1976; Brody 1980).

Public policy arguments.—Others have tied the rise of private insurance to the broad public policy stance of the United States and to specific laws and policies. Jill Quadagno (1984, 1988) links the rise of private pensions to the weakness of public protections and to the persisting political power of industrialists who opposed social insurance. Paul Starr (1982) suggests that the growth of private health insurance was contingent on the failure of a series of public health insurance bills in the early forties.

More specifically it has been argued, as early as 1953 by Louise Ilse and as recently as 1988 by Jill Quadagno and by Beth Stevens, that particular wartime federal policies provided the decisive push to employment-related health and pension insurance. The wartime rise in fringe benefits is thought to be the result of the confluence of three policies: the excess-profits tax, the wage freeze, and the tax-exempt status of

pension and health insurance payments. That is, the wartime excess-profits tax spurred firms to increase their before-tax expenditures, and increased labor turnover encouraged them to do so by raising wages. However, the wage freeze stymied them. Caught in a bind between the excess-profits tax and the wage freeze, employers gained relief in the form of federal rulings that employer payments for pension and health insurance were not covered by the wage freeze and were tax-deductible (Munts 1967; Macaulay 1959). The huge increases in health and pension insurance during the early forties are attributed by many to this set of circumstances, which purportedly caused employers to install or expand fringe-benefit packages in lieu of increasing wages.

The empirical sections below look at welfare work and insurance practices in the early years of the Depression, in the last half of the thirties, and during World War II. Industrial arguments receive little support in the data on the Depression years, because both welfare work and insurance rose during the early 1930s when all of the industrial factors associated with a tight labor market were mitigated by the economic collapse. Moreover, interindustry data from the forties suggest that industrial factors cannot explain wartime increases in insurance coverage. Public policy arguments also receive little support, because the wartime rises in health and pension benefits *preceded* the war-related wage freeze and excess-profits tax.

 Institutional context and interest group goals.—I develop an alternative argument that draws on insights from neoinstitutional organizational theory and institutional theories of public policy, but which fills an important gap in those theories. These perspectives point to the ways in which institutional structures constrain the policy choices made by organizations and by nation-states. At the organizational level, practices and structures become institutionalized in the environment—often in response to public policy inducements—and then diffuse across all sorts of organizations (Meyer and Rowan 1978; DiMaggio and Powell 1983; Tolbert and Zucker 1983). In the process of institutionalization, practices are socially constructed as rational and enter the body of accepted corporate practice. For the most part, organizational theorists have neglected the role of interest groups in promoting salient public policies at the political level, and in helping to institutionalize certain organizational practices (Perrow 1986; DiMaggio and Powell 1991). Some notable exceptions include Fligstein's (1990) work linking federal antitrust policies to the goals of corporate managers and, in turn, to corporate strategy, and Baron, Dobbin, and Jennings's (1986) work linking wartime federal controls on labor turnover to the goals of personnel professionals and, in turn, to changing personnel practices. This study extends this line of research by examining

how public policy shifts changed the organizational and political goals of salient interest groups, and how those altered goals in turn stimulated the growth of fringe benefits.

The new institutionalism in political science has, unfortunately, helped to validate organizational studies that neglect conflict, because as a polemic against the conflict approach to policy-making it has deemphasized interest group behavior (Krasner 1984; Zysman 1983; Skowronek 1982). Some have examined how institutions offer differential resources to different groups (Weir and Skocpol 1983), but few have examined the larger question of how institutional context shapes group goals in the first place and even determines what kinds of groups will emerge (but see Hall 1986). The evidence presented below suggests that institutional theorists should treat group behavior as an intermediate variable between institutional context and organizational and political choices. If the goals of interest groups such as unions vary considerably across time and space, and they do, then institutionalists should be examining the contextual factors that influence those goals.

In the case of fringe benefits, by tracing the effects of public policy on labor and business group preferences and on the insurance industry, we can understand why those groups pursued particular goals at particular points in time. The prevalence of fringe-benefit programs increased in the second quarter of this century largely because public policy created incentives that caused each of these groups to promote health and pension coverage in organizations. This approach highlights a central weakness in the conflict approach, namely the premise that interest groups have predictable, time-invariant goals. In fact, union and business goals switched back and forth between employer-provided insurance and social insurance as public policy changed. The approach outlined here brings the insights of institutional theory and conflict theory together.

How did public policy influence the goals of business and labor groups? First, the combination of early industrial development and late state development left the United States with a large number of industrial workers in the last quarter of the 19th century and with a state that showed no signs of being able to provide them with social insurance. To fill the gap, benevolent societies (later unions) appeared in dangerous industries such as railroads and mining; firms installed their own informal protections for workers; and a private life insurance industry arose to protect families against destitution in the event of the death of the breadwinner (Zelizer 1979). These private forms of coverage would have lasting effects. First of all, the United States developed a strong insurance industry that would promote private insurance and would lobby against public coverage. Second, American Federation of Labor (AFL) leadership came to believe that unions' benevolent functions increased member

loyalty, thus they supported private forms of coverage until about 1932. Then, after the passage of the Social Security Act in 1935, case law surrounding the Wagner Act caused unions to spend their energies fighting for the right to bargain over fringe benefits, which turned them into supporters of private coverage during the late thirties and forties. Third, some business leaders supported social insurance in the teens, but in the absence of public protections the business community at large became wedded to employer-provided insurance by about 1920; they too had come to believe that it helped them win worker loyalty. After the owners of some center firms supported Social Security in the midthirties, the majority of capitalists backed employer-provided benefits during the forties, as part of the fight with unions to keep fringe benefits off the bargaining table. After public policy came to favor unions in benefit negotiations, business leaders led a successful fight for the expansion of Social Security benefits in the early fifties. Public policy context, then, shaped the changing goals of these groups, and in the process it contributed to the rise of private fringe-benefit programs. Moreover, public policies frequently had unintended effects, as when Social Security legislation effected increases in the popularity of private pensions.

THE DATA

This article tries to sort out the origins of employment-related pension and health insurance by examining interindustry data that were collected by the National Industrial Conference Board (NICB) between 1928 and 1946. The data facilitate interindustry and over-time comparisons that make it possible to evaluate competing theories of the rise of insurance. The NICB is a business association that conducts studies of current business practices. It carried out extensive industrial surveys in 1928, 1935, 1939, and 1946 that included questions about informal pensions, mutual benefit associations, pension insurance, and health insurance (NICB 1929, 1936, 1940a, 1947).

   The NICB's industrial relations surveys contain the best data available on the prevalence of early corporate fringe-benefit programs. Earlier studies that employed these surveys to chart the growth of personnel administration and internal labor markets have identified several advantages (Baron et al. 1986; Baron, Jennings, and Dobbin 1988). First, the surveys contain data on thousands of publicly held firms from every region of the country and every industrial sector, and they cover a crucial period of nearly 20 years. In the words of the NICB, "The Conference Board's compilation has remained the only survey of nationwide scope, embracing a wide range of practices, and presenting information in detail" (NICB 1954, p. 1). The board's 1928 survey included 6,085 firms;

the 1935 survey included 2,452 firms with a total of 4.5 million workers, or 15.5% of the national labor force in covered industries (NICB 1936, p. 5); the 1939 survey included 2,700 firms with a total of 5 million employees (NICB 1939); and the 1946 survey covered 3,498 firms with an unreported number of employees (NICB 1947).[3] Second, the surveys are comparable over time because the NICB used consistent survey techniques. It compiled a list of firms from the publications of Dun and Bradstreet, the New York Stock Exchange, and Standard and Poor, and sent surveys to all firms on the list.[4] But in 1928 it was decided to extend "the study in order to develop the picture of industrial relations activities in . . . smaller establishments by oversampling firms with fewer than 250 employees" (NICB 1929, p. v). The 1928 sample includes roughly four times as many small firms as the later samples, and because there is no record of how the NICB increased the participation of small firms it is not prudent to think of the figures for small firms in 1928 as comparable with those for later years.

With the exception of the extra group of small firms in the 1928 survey, it appears that many of the same firms participated in each survey because NICB affiliates regularly took part in its studies (NICB 1954) and because the board used the same procedures to contact firms for each survey. Characteristics of the data also suggest comparability. Average firm size increased less than 1% between the 1935 and 1939 surveys, although average employment was not published in the 1946 report. Moreover the proportion of firms in each industrial category was quite stable over time, and the total number of large firms was stable over the first three waves, which suggests that many firms participated in every survey.

A third advantage is that, for the last three panels, the published reports provide data on some 25 detailed industry categories. This facilitates interindustry and over-time comparisons.

The board's reports also have several disadvantages. First, for 1928, industry-level data were reported only for small firms and only for 10 broad industry categories, thus meaningful interindustry comparisons

---

[3] While the number of firms with fewer than 250 employees was virtually the same in 1939 and 1946 (861 and 867, respectively) there is a substantial increase in the number of large firms. It is likely that more firms fitting the NICB survey profile responded in this year because there was a great deal of interest in the changes in personnel practices wrought by the war economy. Many of the NICB's regular respondents grew dramatically in size during the war, which probably explains the disproportionate increase in responses from large firms.

[4] E. Kay Worrell, Survey Research Center Manager at the NICB, kindly looked into the survey procedures used in the early NICB personnel studies and provided me with this information.

cannot be made between 1928 and later years.[5] Second, the reports certainly exaggerate the prevalence of fringe benefits in the American economy because the surveys were biased toward large publicly held firms—mean firm size was 1,836 and 1,847 employees in 1935 and 1939, respectively. Yet the board saw this size bias as an asset, for "large companies frequently are leaders in inaugurating techniques" (NICB 1954, p. 2). Likewise for the present purposes, the size bias may be seen as an asset because it highlights industry-level trends. A third disadvantage is that firms were not asked how many of their employees were eligible for benefit programs, and we know from the 1946 tables (NICB 1947) that a number of firms offered pensions only to managers. In short, while the surveys surely magnify the aggregate use of pension and health benefits in American firms, they do reflect what was going on in large firms in each sector and they have the asset of over-time comparability.

THE EARLY DEPRESSION

Welfare Work

What caused employers to install welfare practices, which served as precedents for insurance, and what caused them to abandon welfare practices and open the way for insured benefits? Unfortunately, the NICB data pick up welfare work in the middle of its life cycle, which makes sorting out causes somewhat problematic; however, industries that are outliers in the 1935 survey do provide some useful insights about industrial causes.[6]

The industries that show exceptionally high scores on both informal pensions and mutual benefit associations are the utilities (gas and electricity) and trade (see table 1). First, the utilities were both capital and skill intensive, but only a sixth of the firms surveyed were unionized. The other blue-collar nonmanufacturing industries—transportation/communication and mining—were likewise capital and skill intensive and shared high scores on welfarism, but they were highly unionized (50% of the capital-intensive firms and 57% of the skill-intensive firms were unionized; see NICB 1936). This suggests that early welfare work was driven by efforts to retain skilled employees in capital-intensive sectors rather than by unionism per se. Table 2, which reports every pension program that the NICB could locate in 1925 by industry, confirms this:

[5] While the NICB surveyed thousands of organizations, they only reported industry-level figures and they disposed of the original questionnaires.

[6] Because the 1928 study reported interindustry data only for small firms, the 1935 data give a better picture of differences across industries. Note that the marginals for large firms changed little between these two panels, which suggests that most growth had occurred by 1928 and that the decline of these practices had not yet begun.

TABLE 1
PREVALENCE OF BENEFITS BY INDUSTRY AND YEAR (%)

Industry	Informal or Individual Pensions				Mutual Benefit Associations				Group Pensions				Health/Accident Insurance			
	1928	1935	1939	1946	1928	1935	1939	1946	1928	1935	1939	1946	1928	1935	1939	1946
Agricultural implements		10.0	0			50.0	41.7			20.0	25.0	31.5		10.0	20.8	
Automobiles and parts		21.9	(15.4)	10.9		31.5	(21.2)	16.3		0	(1.0)	48.7		53.4	(53.8)	78.3
Aircraft and parts	15.1			5.1	4.8		18.0	10.3	4	14.0	20.3	60.8				87.2
Chemicals and chemical products		31.6	40.6	11.6		26.3	23.1	7.2		7.8	12.8	40.7	9.5	34.2	43.8	67.4
Electrical manufacturing		23.3	18.8	7.7		28.9	25.7	8.8		15.0	16.7	54.0		28.9	40.2	60.8
Food products	4.8	33.6	29.5	14.6	.2	25.7	17.9	8.1	0	2.7	4.7	25.9	8.7	33.6	31.4	61.1
Leather and its products	2.7	21.3	14.0	15.5	4.8	24.0	26.6	13.8	.0		.9	3.3	8.0	30.7	28.1	70.7
Lumber and its products	4.0	12.3	14.0		5.6	16.0	17.4		.0				9.4	26.4	40.5	
Building materials and supplies				9.3				11.2				31.8				95.3
Machines and machine tools		18.6	23.8	9.5	5.7	37.1	34.2	12.1	2	5.7	5.0	39.1		25.8	29.2	69.6
Metals and metal products	3.9	21.0	22.6	11.3	4.5	30.6	23.6	10.5	0	6.5	6.7	31.8	11.7	33.5	38.1	64.3
Paper and its products	6.7	32.3	35.8	14.3		31.0	24.6	14.8	0	5.2	4.5	37.1	15.2	49.0	44.0	75.7
Petroleum and its products		13.5	21.3	3.8		35.1	27.7	15.4		29.7	51.1	73.1		29.7	42.6	73.1
Printing and publishing	4.9	22.1	29.6	15.6	5.8	27.9	24.5	18.9	1.3	9.3	13.3	35.6	13.3	27.9	34.7	63.3
Rubber	10.7	12.5	21.1	10.2	3.6	30.0	15.8	6.8	0	7.5	5.3	39.0	14.3	47.5	52.6	72.9
Stone, clay, and glass		21.2	22.0			4.5	14.0			1.5	9.0	40.7		31.8	31.0	
Glass																
Textiles and clothing	3.2	22.8	23.1	15.2	2.2	21.3	13.4	6.8	2	5.5	5.6	28.0	11.9	18.1	23.5	55.6
Manufacturing industries	4.7	22.7	24.5	11.6	4.4	27.6	22.5	10.4	2	7.0	8.8	37.6	10.9	31.2	35.6	66.9
Banking		15.0	22.6	7.5		2.5	4.8	0		37.5	51.6	78.5		10.0	12.9	13.9
Insurance		15.4	33.3	11.1		7.7	7.1	3.7		46.2	59.5	75.9		30.8	23.8	42.6
Gas and electricity		37.7	38.5	5.8		42.0	31.9	19.8		40.6	50.5	84.3		34.8	38.5	44.6
Transportation and communication	(9.5	24.3	16.7	14.3 (10.7))	(9.5		32.4	24.4)	(0	48.6	38.5	61.2 (66.6))	(14.3	40.5	34.6	38.8 (48.0))
Transportation			3.8	17.9				30.6				76.9				65.4
Communication and broadcasting	1.0	42.1	40.2				3.8	21.3	1.0	11.3		53.8	2.0	21.1	18.6	39.6
Wholesale and retail trade		32.1	23.8		4.1	51.3	32.0	23.6		6.6	4.8			10.7	31.0	
Mining				18.5		28.6	7.1				7.1					
Coal and coke							11.1	14.6				29.6				63.0
Nonmanufacturing industries	3.8	29.4	29.9	11.3	6.2	32.4	21.1	14.6	8	30.9	34.3	67.4	6.2	27.6	27.0	38.9
All industries (firms with fewer than 250 employees)	4.6	15.9	14.6	7.4	4.5	16.8	10.8	3.9	2	3.9	6.0	24.8	11.0	24.5	25.7	53.3
All industries (firms with 250 or more employees)	26.4	27.4	30.5	12.7	29.7	34.0	27.9	13.3	1.9	13.4	16.1	47.5	15.5	33.8	38.3	67.2

NOTE.—Industry-specific and sectoral data for 1928 pertain only to firms with 250 or fewer employees Several industries have been collapsed to facilitate comparisons over time textiles and clothing are combined, "chemicals" includes drugs, dyes, paints, pigments, varnishes, soap, and toilet preparations; and "metals and metal products" Includes iron and steel and nonmachine metal products For some other industries that could not be collapsed to match over time, percentages are reported within parentheses that span several categories in problematic years (e.g , transportation, communication, and broadcasting) Marginals by sector and size include some industrial categories not listed here

TABLE 2

Number of Pension Plans in Operation by Industry, 1925

		Noncontributory		
	Contributory (Insured)	Discretionary	Limited Contractual	Total
Chemicals	1	5	4	10
Food	3	3	0	6
Metals	5	38	17	60
Paper and printing	0	8	1	9
Petroleum	0	3	8	11
Textiles	1	6	0	7
Miscellaneous	2	3	4	9
Banking	9	8	0	17
Insurance	0	8	0	8
Mining	1	4	1	6
Railroads	1	36	0	37
Trade	2	9	1	12
Utilities	2	37	7	46
Total	27	168	43	238

Source NICB (1925, p 15)

informal pensions were common in the highly unionized railway industry, but they were also common in the largely nonunion metals and utilities sectors (Wolman 1936). The relationship that some have found between unionism and welfare practices, then, may be spurious. Industries that were both capital and skill intensive could ill afford labor discontent, which made it difficult for them to fight unions and also spurred them to install welfare practices to placate workers.

Retail and wholesale firms were the most likely to use both welfare practices. Carter and Carter (1985, p. 590) suggest that large retail stores adopted welfare work practices because their profitability depended on presenting an affable, clean, and healthy face to the public (see also Nelson 1975; Labor Statistics Bureau 1917). By offering pension and health coverage, firms kept ill and superannuated employees at home. At first glance that thesis seems to be challenged by the fact that in white-collar banking and insurance firms, which likewise depended on customer relations, informal pension plans were rare and mutual benefit associations were unheard of. Table 1 shows that those industries used formal pension insurance instead of informal pensions, yet for health insurance, banks reported the lowest use of any industry and the insurance industry itself reported only average use. Why did banking and insurance differ from trade in their use of health coverage? Contemporary

studies found that, instead of offering health coverage, some one-half to two-thirds of banks and insurance companies guaranteed salary maintenance for injured or ill workers, but this practice was unworkable in wage-based industries where the time clock reigned (Baker 1940; NICB 1937).

Most analysts date the decline of welfare work to the Depression, arguing that once problems of turnover and union activism subsided, employers could ill afford to be charitable, and that Roosevelt's hardline stance against industrial paternalism discouraged welfarism (Brandes 1976; Brody 1980; Achenbaum 1986). Yet figure 1 shows no decline; on the contrary, both practices showed small increases among large firms (see table 1 above). Moreover, the 1935 study reports that only 4% of those firms that had used informal pensions had discontinued them and only 8% of those that had used mutual benefit associations had discontinued those (NICB 1936, p. 11; see also Parks 1936). Some health and pension welfare programs may have expired when firms went bankrupt and closed their doors, but few of the firms that endured canceled their programs. The 1939 figures show that, even after the Wagner Act had expanded union membership, firms did not abandon welfare practices.

Why, then, do analysts date the end of welfare practices to the Depression or to the Wagner Act? Sanford Jacoby (1984, p. 41) suggests that it is because paternalistic "old welfare work" practices such as housing programs, thrift plans (e.g., stock purchase programs through payroll deductions), recreational programs, and educational programs were discontinued, but "new welfare work" practices, which tied benefits to loyalty and job tenure (e.g., pensions and mutual benefit associations) actually grew. Indeed, the NICB found that among large firms home purchase plans declined by 48% between 1928 and 1935, company housing programs dropped by 28%, and stock purchase plans fell by 56% (NICB 1929, 1936). But firms did not cancel health and pension forms of welfare work.

In sum, the argument that employers installed pension and health care forms of welfarism to quell turnover among skilled employees receives some support in the NICB data. Yet the evidence for arguments about labor turnover and union busting is mixed, because firms did not abandon these forms of welfarism either in the early thirties when labor turnover was no longer a problem, or in the late thirties after their employees had joined unions.

Health and Pension Insurance

The prevalence of employment-related health and pension insurance increased significantly between 1928 and 1935 (see table 1 and fig. 1).
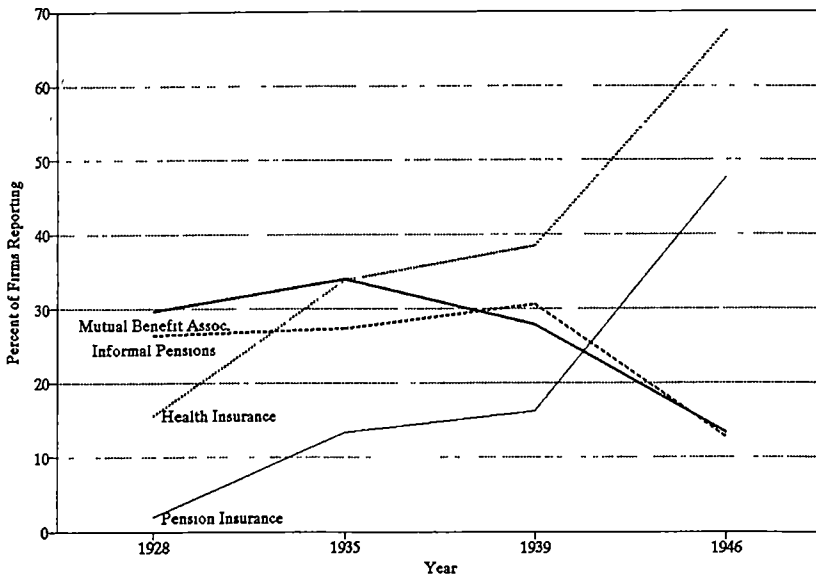
FIG. 1.—Pension and health benefits among large firms, 1928–46

Among the large firms surveyed, health/accident insurance more than doubled, from 16% to 34%, and group pension plans rose from under 2% to over 13%. In addition, existing pension and health insurance programs were seldom abandoned despite the unsteady economy. In its 1935 study the NICB found that only 7% of all pension and health insurance programs had been canceled since the onset of the Depression (NICB 1936, p. 11; see also Parks 1936). These trends belie every industrial explanation of the origins of insurance, because the Depression mitigated the problems associated with turnover, union activism, organizational growth, and labor-market segmentation. However, public policy had marked effects on the aims of the insurance industry, labor unions, and business groups, all of whom successfully promoted private insurance at the organizational level and provided weak support for social insurance.

Public Policy and Interest Group Goals

Three important groups promoted private health and pension insurance during the twenties and early thirties; their support can be traced, broadly, to institutional context and, in a more specific sense, to public policy. The industrial revolution brought huge firms and enormous unions to the United States in the latter decades of the 19th century, but the American state remained singularly weak and disorganized. It was

clear that the state would be unwilling, and unable, to provide adequate social insurance coverage for industrial workers, and unions and businesses soon stepped in to fill the gap by providing coverage of their own. American Express introduced pension coverage in 1875, and by 1925 over 200 large firms were offering pensions (NICB 1925, 1929; National Personnel Association 1922, p. 6). Health and accident coverage was first offered in the dangerous mining and railway industries by benevolent societies that soon evolved into unions (Brandes 1976). Business and labor groups that had developed their own forms of insurance opposed social insurance, such as the state health programs advocated by Progressive reformers in the 1910s, so that by the mid-thirties only private health and pension coverage existed (Anderson 1985, pp. 66–73). Meanwhile, in the 1800s a large life insurance industry, which surpassed European insurance industries in size and sophistication, grew up alongside new industrial enterprises with the encouragement of state legislatures (Zelizer 1979). By the early 1930s, their experiences with private insurance turned insurers, unions, and business leaders into effective supporters of private coverage.

The insurance industry strategy.—The presence of a large insurance industry, with sophisticated actuarial techniques, that was poised to expand insurance coverage during the Depression contributed to the growth of private health and pension insurance. The Depression proved to be an unexpected boon to the insurance industry, whose assets increased by one-third between 1929 and 1935, in part because of the increased economic insecurity during that period (James 1947, p. 293). While large insurers saw declines in their commercial business, they also saw substantial increases in group insurance plans for employees. The number of companies writing health and accident insurance rose by 20% between 1929 and 1931 alone (Insurance Almanac 1930, p. 1189; 1932, p. 1016). For 1931, the eight largest group insurers reported that one-third of their premium income came from group insurance *other than life insurance,* such as health and pension insurance, compared with only 1% from this source in 1926 (Bureau of Labor Statistics 1932–43 [1932, p. 53]).

The increasing popularity of income-protection insurance was partly a result of the new sales strategies insurers adopted during the Depression. Insurers encouraged salesmen to cast health and accident coverage as "income insurance" that would guard families against destitution (Tiger 1932, p. 11; Landers 1935, p. 13). In response to a 250% increase in disability claims against life policies during the early Depression, the major insurance firms separated life and disability coverage and began offering these popular forms of coverage in low-cost package deals that included pension and health coverage as well (NICB 1934, p. 36; Bureau

of Labor Statistics 1932–43 [1933, p. 54]; Stone 1957, p. 154). Insurance companies also encouraged employers to sustain coverage for employees by (*a*) using "work sharing" instead of layoffs, which helped to keep employees on the insurance rolls, and (*b*) continuing to pay premiums during temporary plant shutdowns to prevent policy cancellation (Bureau of Labor Statistics 1932–43 [1933, p. 53]). Among insurers, industry publications encouraged firms to try to replace their lost commercial business with these forms of coverage. By 1935 an industry trade journal suggested that a "'well balanced' agency should have at least a third of its casualty premiums in the health and accident line" (*Rough Notes* 1935, p. 11).

While they were promoting private insurance at the firm level, insurance industry leaders were lobbying Congress for Social Security legislation that would favor the industry. They promoted an amendment that would exempt employers with private insurance schemes from Social Security taxes, with the goal of shaping legislation that would cause most employers to adopt private insurance. H. Walters Forster, a partner in a Philadelphia agency specializing in group annuity coverage, organized a massive campaign among insurance industry executives and employers in support of this amendment, but President Roosevelt insisted on public coverage, and in the Senate committee the final vote was a tie, which kept the amendment out of the act (Witte 1962, p. 161).

Union goals.—On the other hand, AFL leadership, under Samuel Gompers, favored private benefits programs consistently until 1932, even though some AFL locals had come to support social insurance instead. Gompers's position can be traced to two factors. First, he preferred union-provided insurance because he believed that union health benefits won members. His own union career had begun with the cigar makers in New York, where he proposed union sickness and death benefits in the late 1870s and saw a tenfold increase in membership (Starr 1982, p. 249). Railway unions, which had their roots in benevolent societies formed to protect members against calamity, shared Gompers's sentiments. Of course, those union benefit programs had appeared in the first place to make up for the absence of public social insurance.

Second, Gompers and other AFL leaders were suspicious of proposals for public coverage because federal policy in the teens and twenties had been so firmly antiunion. Thus in 1916 Gompers not only failed to lobby for social insurance legislation, he testified against it before the House Committee on Labor (Anderson 1985). The legacy of early union benefit programs and the federal government's antiunion stance was an American union movement led by men who opposed social insurance and favored private protection (Achenbaum 1986, p. 84; Witte 1962).

Business preferences.—Business leaders had briefly advocated social insurance in the teens, but they soon switched their support back to private coverage. European business leaders in skilled, highly paid industries often supported social insurance because it shifted responsibility for coverage to the state, so it is somewhat surprising that their American counterparts largely backed private coverage.

The record suggests that high-wage employers backed private insurance schemes for the same reason Gompers did; they believed they had won worker allegiance by offering benefits (Quadagno 1984, p. 636). Firms at the forefront of welfare capitalism, such as Western Electric, had first installed welfare programs to win the hearts and minds of employees and had apparently achieved success. Consequently, such business organizations as the National Association of Manufacturers, the Chamber of Commerce, and the National Civic Federation withdrew their support for public health coverage during the teens when members argued that employer-provided coverage served important labor control functions (Starr 1982, pp. 250–52; Anderson 1985, p. 86). That was the turning point for business support of state health insurance. In the 1920s, business associations (NICB 1925) and government agencies (Bureau of Labor Statistics 1928) promoted the idea that fringe benefits could increase productivity and help control labor dissent (Berkowitz and McQuaid 1980, pp. 82–84).

In the 1920s, industrial relations professionals, echoing welfare capitalists, argued that health insurance was an effective antiunion device because it demonstrated the goodwill of the employer and circumvented the union bargaining process. Auto industry firms, for instance, had introduced health coverage to undermine the efforts of the United Auto Workers, an organization adamantly opposed to unilateral fringe-benefit programs because they interfered with the union's role of representing the workers (Munts 1967, p. 9). General Motors introduced a group health insurance plan in 1926, finding that when unions objected to the plan they alienated prospective members (Munts 1967, p. 48). This strategy of using benefits to quell unionism was widely promoted in the contemporary literature on union avoidance. Accordingly, in table 1 we find that industries that faced the most active union organizing efforts—autos, paper, and rubber—show remarkable rates of private health coverage (NICB 1940a; Pelling 1960). As early as 1928, the paper and rubber industries reported unusually high rates of health insurance coverage, and by 1935 autos, paper, and rubber reported rates that were nearly double the average and at least 13 percentage points higher than the next highest manufacturing industry (see table 1).

Recent policy shifts and business preferences.—Three contemporary

policy changes encouraged firms to adopt private health and pension insurance. First, tax code changes in 1926 made corporate contributions to insured pensions deductible, and after the corporate tax increases of the early thirties this made insured pensions increasingly attractive (Schieber 1982; Stevens 1988). Indeed, between 1925 and 1935, insured pensions rose from 11% to 30% of all pensions (NICB 1925, 1936). Second, Roosevelt's New Dealers tried to put an end to welfare work first by prohibiting employers from requiring participation in "company unions," which were at the core of welfarism, in the National Industrial Recovery Act of 1932, and then by outlawing "company unions" altogether in the Wagner Act of July 1935 (Schlesinger 1958). Stuart Brandes argues that, on top of the economic exigencies of the Depression, this attack on company unions brought down "the whole array of welfare practices" (1976, p. 144). While we know that welfare work survived the early thirties, the antiwelfare stance of the New Dealers undoubtedly encouraged employers to choose health and pension insurance over uninsured plans (see fig. 1). In addition, the banking and insurance industries faced special federal incentives during these years. Expanded Depression-era regulation of financial institutions through the Securities and Exchange Commission and the Federal Deposit Insurance Corporation made banks and insurance companies accountable for their outlays. Because firms had difficulty justifying to federal regulators their informal, discretionary pension payments, many switched to insured pension schemes (Schlesinger 1958; Romasco 1983). Thus by 1935, 38% of all banks and 46% of all insurance companies reported offering pension insurance (NICB 1936). These are among the highest rates of use for the time. Health insurance plans doubtless remained rare in those sectors because of the practice of salary continuation, which did not raise the eyebrows of auditors.

In sum, Depression-era industrial conditions do not seem to have put an end to welfare work or retarded the growth of employment-related insurance. Instead, American public policy, specifically the paucity of public income protections, had created a sophisticated insurance industry, ready to promote income-loss insurance during the Depression, and had led business and labor leaders to develop their own private forms of illness and pension coverage. Privately organized insurance schemes gained constituencies among union and business leaders, as Philip Selznick (1948) would predict, and those leaders came to prefer private forms of coverage to public forms in the twenties and early thirties. Finally, specific public policies encouraged firms to adopt insured benefits and to abandon informal benefit programs. As a result private insurance rose during the early years of the Depression.

THE LATE THIRTIES

The Social Security Act

Policy analysts anticipated that Social Security and the Wagner Act would have marked effects on employment-related insurance. Social Security was expected to render private pension insurance obsolete, at least for nonmanagerial employees. By contrast, the Wagner Act led to a rise in unionism that was expected to result in the growth of all sorts of employment benefits (NICB 1936, 1940a). Yet table 1 shows almost no change in the aggregate figures for health and pension insurance between 1935 and 1939. Was insurance really stagnant over these years? A closer examination of the data illuminates some unanticipated consequences of the Social Security Act. The section on the war years, below, deals with the specific effects of the Wagner Act.

The NICB surveyed several hundred firms in 1939 to ascertain the effects of Social Security legislation on private pension schemes. It found that less than one-tenth of those offering private pensions in 1935 had canceled them by 1939. "This delay in making necessary adjustments may be explained by the constant agitation for certain fundamental changes in the law which began almost as soon as it became effective" (NICB 1939, p. 24). Roosevelt argued for the expansion of the Social Security Act while his opponents tried to get it declared unconstitutional in the courts; firms waited to see who would prevail.

However, the NICB data suggest that, in high-wage industries, the low Social Security benefit levels caused substantial numbers of firms to purchase private supplementary insurance without delay. The maximum benefits to be paid out by the old-age insurance program of Social Security were set below the minimum wage and were compressed. An employee who was earning $100 a month and who was fully vested would receive $35 a month upon retirement, or 35% of her regular wages. Yet an employee earning $250 a month would receive only $56 a month upon retirement, or 22% of her regular wages. Because benefits were limited to the first $3,000 of wages, an employee earning $5,000 annually would also receive $56 monthly, or 13% of her regular pay (Ilse 1953, p. 297). Consequently, the NICB data show the greatest percentage gains in the use of insured pensions in industries with highly paid employees (Bureau of Economic Analysis 1986; NICB 1939). In the manufacturing sector, chemicals, electrical manufacturing, petroleum, and printing saw the greatest percentage increases in group pensions between 1935 and 1939[7]—gains of at least 40%. According to the 1940 census, these four

[7] This is true with the exception of three industries that saw large *percentage* increases because less than 3% of firms had pensions in 1935.

industries had at least twice as many professional and semiprofessional workers (8.6%–10.2%) as other manufacturing industries (0.6%–3.8%), with the exception of machines (10.2%; Bureau of the Census 1940).[8] These industries also reported wages that were 25%–45% above the national average in 1939 (Bureau of Economic Analysis 1986, p. 279). Many firms had held off adopting private pension schemes because they had anticipated federal coverage, but the passage of this weak public coverage evidently convinced those in high-wage sectors to wait no longer (NICB 1939; Ilse 1953). The effects of weak Social Security coverage can also be seen among banks, whose employees were excluded in the 1935 legislation. As a result, we see in table 1 that private pension coverage in banking increased 14 percentage points between 1935 and 1939. In short, in sectors where the Social Security Act provided the weakest pension coverage, passage spurred firms to adopt private plans. The architects of Social Security legislation had not anticipated that outcome. By 1939 the act had caused twice as many firms to install private pension programs as to cancel existing plans (NICB 1939, p. 24). Helen Baker, of Princeton's Industrial Relations Section, wrote in 1940 that the act had "encouraged rather than retarded voluntary action [private pension coverage] by the employer" (1940, p. 10).

The act also caused two-thirds of the firms that offered pension insurance in 1935 to replace it with supplemental insurance designed to dovetail with public coverage by 1939 (NICB 1939). Most of the new programs (66%) were graduated, so that private pension coverage increased with income as a function of the decline in public coverage, which meant that employees would receive total retirement benefits equivalent to a flat percentage of their working income (NICB 1940b).

Social Security legislation also spurred firms with informal pension plans to update them, and in the process many switched to the contemporary insured (and tax-deductible) form of pension coverage. The NICB found that one-fourth of firms that used the old-style self-administered (informal) plans in 1935 had already switched to (formal) annuity plans by 1939, yet no firms had made the reverse transition. In addition, 34% of operating group annuity plans had been changed from the older employer-paid form to joint contributions (nonrevocable), and only one company had made the reverse change (NICB 1939, pp. 25–26). The Amended Social Security Act of 1939 apparently reinforced this trend, as employers who were revising pension practices opted for the modern insured form (see table 1).

---

[8] The figure for agricultural implements was not published separately. The change in stone, clay, and glass probably represents a substantial change in glass, which is highly skilled, and less of a change in stone and clay, which are not.

In sum, the passage of Social Security legislation had an unexpected positive effect on the prevalence of private pension insurance in selected sectors. Most firms delayed canceling private pension insurance, high-wage firms installed supplementary plans, banks installed private pensions, and most firms that had pension plans replaced them with special supplemental plans, and, in the process, many switched from informal to insured pension plans.

THE WAR YEARS

The growth in group health and pension plans during the first half of the forties was astounding. Analysts have, quite naturally, pointed to effects of the war to explain these increases, but, in the light of the available historical evidence and the NICB data, current thinking about what went on during the war warrants several revisions. First, the interindustry data show little effect of turnover, labor-market segmentation, or increases in the size of organizations on the growth of insurance between 1939 and 1946. Insurance became more popular in every industry. Second, when it comes to public policy, the effects of wartime tax changes and wage controls have probably been overstated, yet the 1939 Social Security revisions had marked effects on the incidence of private pension programs. Third, pension and health insurance rose at a rapid rate in unionized sectors, but not because unions won benefits in their negotiations with employers. Instead, public policy, in the form of case law surrounding the Wagner Act, spurred employers to offer insurance unilaterally to subvert unionism; at the same time this caused unions to channel their energies toward winning the right to bargain over fringe benefits.

Industrial Arguments

Turnover.—While the economywide labor turnover rate increased dramatically during the war, from 3.2 monthly separations per 100 workers in September 1940 to a peak of 6.3 in September 1943, available interindustry data for the manufacturing sector in 1943 show no relationship with the prevalence of fringe benefits in 1946. The Bureau of Labor Statistics reported turnover rates for nine separate industries in 1943, at the peak of labor turnover, and the industries that had the highest turnover levels—shipbuilding, glass, cotton textiles, and clothing—showed average or below average levels of pension and health insurance in the NICB's 1946 study (the low turnover industries were aircraft, autos, metals, electrical machinery, and machinery; Bureau of Labor Statistics 1932–43 [1943, p. 1241]). Further, if benefits were indeed adopted to prevent wartime turnover we would expect munitions firms to have been

the leaders, yet the shipbuilding, aircraft, explosives, and converted automobile industries reported overall pension and health insurance rates of 37% and 77%, respectively. These rates are not substantially different from the aggregate figures for manufacturing, in which 38% of all manufacturing concerns offered pensions and 68% provided health insurance. Health insurance was slightly more prevalent there, but it had always been popular in these dangerous heavy manufacturing sectors. For instance, the automobile industry had the highest incidence of health insurance even in 1935.

The more diffuse effects of turnover throughout the economy are difficult to gauge; however, several factors suggest that employment stabilization was not the principal driving force behind the adoption of pension and health insurance. First, the war industries that grew the most (aircraft, shipbuilding, and the converted auto industry) paid well even before the war and did not have to devise new strategies to lure workers (Bureau of Economic Analysis 1986). Second, the federal government exercised a variety of controls, through agencies such as the War Production Board and the War Labor Board, that dampened turnover. These boards were most vigilant in munitions (Baron et al. 1986). Finally, as discussed below, the National War Labor Board (NWLB) wage freeze of 1942 was commonly circumvented, which meant that employers did not have to turn to nonwage forms of remuneration to attract workers.

Labor-market segmentation.—Widespread increases in fringe benefits appeared across all industrial sectors, which contradicts the notion that modern labor-market bifurcation had its roots in part in the divergence of benefit plans during this period. No matter how we categorize primary labor-market industries, they do not differ substantially from other manufacturing industries. In the aggregate, 38% of all manufacturing firms reported providing group pensions, and 68% offered health insurance in 1946. Core firms reported 39% and 70%, respectively; capital-intensive firms likewise reported 39% and 70%; and durable goods firms reported 36% and 69%.[9] While the large average firm size in the NICB studies may mean that the data represent core firms even in peripheral industries, other analyses of the NICB data have found marked sectoral differences in labor practices as predicted by labor segmentation theorists (Baron et al. 1986, 1988).

Organizational scale.—The number of employees in a firm appears to

[9] Core industries are aircraft, autos, chemicals, electrical goods, machinery, metals, and rubber (Hodson 1978; Gordon et al 1982) Those with high capital-output ratios are lumber, chemicals, glass, metals, paper, petroleum, and printing (Kuznets 1961, pp. 214–15; Bureau of the Census 1947, p. 159). Durable goods industries are automobiles, aircraft, electrical manufacturing, building materials, machines, metals, stone, clay, and glass (Bureau of Economic Analysis 1986).

be an excellent predictor of pension and health insurance coverage, a fact that supports the organizational maxim that increased scale is associated with increased formalization (Pugh et al. 1969; Blau and Schoenherr 1971). Turning again to table 1, insured group pensions were nearly twice as prevalent in large firms as they were in small firms in 1946 (47.5% vs. 24.8%), and health insurance for large firms was 14 percentage points higher than it was for firms with fewer than 250 people. Yet the data contradict the idea that firms formalize certain practices once they reach a specific size and the corollary that aggregate increases in formalization over time are the direct result of increases in average firm size. In 1946, the percentages of *small* firms reporting pension (24%) and health (53%) coverage were greater than the percentages of *large* firms reporting that coverage in 1939 (16% and 38%, respectively). This suggests that historical factors are more important than size alone, even if large size increases the likelihood that an organization will adopt certain institutionalized practices.

Public Policy Arguments

Interindustry evidence suggests that wartime industrial problems do not account for increases in the prevalence of fringe benefits. How did the wartime wage freeze and tax code changes affect the incidence of private pension and health insurance? I will argue that policies that were intended to popularize insurance did not have that effect, and that policies that were not expected to popularize insurance did.

   The NWLB wage freeze.—Analysts of fringe benefits have frequently argued that the wartime wage freeze and changes in tax policy stimulated private benefit expansion (Munts 1967; Macaulay 1959; Stevens 1988). The outbreak of war in Europe quickened the American economy, increasing the bargaining power of unions and expanding labor conflict. Roosevelt responded by establishing the National War Labor Board in January 1942, the month following the Japanese attack on Pearl Harbor. The board's most momentous decision came that year in the "Little Steel" wage increase case. Bethlehem, Republic, Inland, and Youngstown steelworkers were demanding raises, but steel industry leaders wanted to stabilize wages and Roosevelt was seeking an economywide wage-price freeze. The NWLB decision was a compromise that tied wage increases to inflation, thereby freezing real wages, and the decision served as a precedent for firms throughout the economy (Seidman 1953; Civilian Production Administration 1947; Kerry 1980; Bureau of Labor Statistics 1932–43 [1940–43]). In 1943, the board ruled that pension and insurance benefits were not subject to the freeze. It is widely believed that this

exemption spurred firms to increase benefits in lieu of increasing wages to attract and retain workers.

The Revenue Act of 1942.—The war brought a new tax code, designed in part to prevent war profiteering, which taxed up to 90% of any profits that exceeded prewar levels. This gave corporations a strong incentive to lower their taxable income. One effect of the law was to encourage new capital investment, but a number of analysts suggest that the excess-profits tax also created a compelling reason to expand tax-exempt fringe benefits. Because wage increases were limited by the NWLB freeze, employers could increase their deductible expenditures for labor only by expanding fringe benefits such as pension plans. Moreover, in 1943 the Internal Revenue Service (IRS) extended the tax-exempt status to payments for health insurance, which gave employers another vehicle to reduce pretax profits (Macaulay 1959). Stevens (1988) suggests that the wage freeze and the 1942 Revenue Act thus combined to help popularize fringe benefits. The result was a fivefold increase in employer contributions to pension trusts, from $171 million in 1941 to $857 million in 1945 (see also Munts 1967; Macaulay 1959; Ilse 1953; Quadagno 1988). For both workers seeking increased remuneration and employers trying to compete for personnel in an increasingly tight labor market, fringe benefits now offered an avenue for circumventing the wage freeze.

The record suggests that these policies came too late in the war to have caused the rises in pension and health insurance that we see between 1939 and 1946 (Stevens 1990). The Revenue Act of 1942 did not take effect until 1943; the NWLB did not begin to exempt benefits from the wage freeze until 1943; and the IRS extension of its pension-exemption policy to health insurance only took effect in 1944. While direct U.S. participation in the war effectively began in 1942, industry began gearing up for the war in 1939, and by 1943 employers had already made adjustments to retain workers. Indeed between 1939 and 1943, when new federal policies first took effect, employment rose by 49%; it increased only another 4 percentage points to the wartime peak in 1944 (Bureau of Economic Analysis 1986, p. 275). Annual data on the number of insured American employees show conclusively that the bulk of the increases occurred before federal policy changes first took effect in 1943. The number of employees carrying group pension insurance rose from 720,000 in 1939 to 1.2 million in 1943 and to 1.47 million in 1946 (Ilse 1953, p. 315). The number carrying group health/accident insurance rose from 3.5 million in 1939 to 6.5 million in 1943 and to 7 million in 1946 (Ilse 1953, p. 189).[10]

[10] The figures for 1946 may be slightly deflated relative to those for 1943 because Ilse's source changed in 1945. As a result, between 1944 and 1945, Ilse's figures show a 4%

Moreover the wage freeze probably had little effect even in the last two years of the war because firms found ways to circumvent it. Because raises were allowed in conjunction with promotions, as early as 1943 three out of five firms were paying wage increases for "in-grade progression," or promotions to progressively higher job titles within the same job—a practice that, because it discouraged turnover, was permitted by the NWLB (Jacoby 1985, p. 264). The wage freeze, then, did not prevent employers from raising wages when they wanted to. This practice appears to undermine the wage freeze/tax increase argument because it enabled firms to retain valued workers and at the same time decrease before-tax profits without installing fringe benefits.

Analysts have reasonably assumed that these policy changes had the effects their authors intended. By exempting fringe benefits from the wage freeze Roosevelt's NWLB hoped to encourage the adoption of insurance coverage. The administration had expected that the tax-deductible status it extended to health insurance would foster fringe-benefit programs. The architects of the 1942 Revenue Act included a clause requiring that tax-exempt pension programs cover 70% of employees and prohibiting discrimination against low-wage employees in order to encourage firms to expand private pension coverage.[11] While these policies probably did encourage firms to install insurance programs, they came too late to have caused the massive wartime increases in fringe benefits.

The Amended Social Security Act of 1939 and Business Strategy

The 1939 amendments to Social Security appear to have had an unintended positive effect on private pension programs that has not been documented in the recent literature. The act as amended encouraged employers to favor private insurance plans for several reasons. On the one hand, the 1939 amendments made it clear to employers who were awaiting Roosevelt's promised benefit increases that those increases would not be soon forthcoming. Employers felt compelled to adopt supplementary pension plans for highly paid employees. On the other hand, Social Security provided a foundation retirement wage, which made private pensions relatively cheap. Employers could now try to win the hearts of workers through pension programs without spending much money.

Roosevelt had hopes for an incremental expansion of Social Security

---

decline in pension coverage and a 9% decline in health coverage But this has little effect on the overall picture.

[11] As Louise Ilse (1953) points out, programs that covered only employees who earned over $3,000 were not considered discriminatory

after 1935, and the intent of the 1939 legislation was to liberalize coverage (Achenbaum 1986, p. 26). Thereafter benefits would be calculated on the basis of average income at the time of retirement rather than on total lifetime contributions; this was expected to increase pension income because employees typically reach peak earnings just before retirement. The 1939 law called for employees to be fully insured with as little as two years of participation in the plan, rather than 40 years, and the first benefits became payable in January 1940, rather than in January 1942 (Achenbaum 1986, p. 30). The amendments also called for benefits, but not contributions, to increase with each dependent and for dependent payments to continue after the death of the insured employee.

But federal pension insurance was a zero-sum game because fiscal conservatives in Congress refused to increase total expenditures. Liberalizing entitlements meant that projected benefits per recipient would not rise significantly (Achenbaum 1986, p. 34). The 1939 act also postponed a scheduled increase in contributions that might have permitted benefits to rise with inflation. While the 1939 legislation called for a slight increase in projected benefits, it would be over a decade before Congress approved new increases.

Had benefits been set at a fixed percentage of income, inflation would have had no effect in real dollars. Instead, the benefit structure caused projected payments to retirees to decline dramatically in real terms between 1939 and 1946 as average wages doubled. First, the plan covered only the first $3,000 of income. For a person who earned $3,000 in 1939 and whose income doubled by 1946, the maximum monthly benefit of $56 declined from 22% to 11% of working income. A growing number of Americans crossed the $3,000 threshold between 1935 and 1946, as that figure declined from 2.6 to 1.26 times the average full-time income. Second, those who earned less than $3,000 also lost in terms of projected benefits because the reverse-graduated plan dictated that benefits would decline, as a percentage of income, as income increased (NICB 1939, p. 41; Bureau of Economic Analysis 1986, p. 279). By failing to increase benefits in 1939 to counteract the effects of inflation, and by delaying an increase in contributions, Congress demonstrated that public coverage would not provide an adequate retirement wage.

Contemporary publications predicted a substantial increase in private plans in response to the 1939 amendments, and the data show such an increase (NICB 1939; 1940b). First, there is a remarkable aggregate increase; table 1 shows that between 1939 and 1946 the incidence of firms offering pensions rose from 7% to 38% in manufacturing and from 31% to 67% in nonmanufacturing. Second, data presented in table 3 suggest that there were large increases in supplemental pension plans for the managerial employees for whom Social Security benefits would replace

TABLE 3

PERCENTAGE OF FIRMS REPORTING GROUP PENSIONS BY TYPE OF PLAN, 1946

	FIRMS REPORTING FORMAL PLANS			
	For Wage Earners	For Salaried Employees	For Employees Earning over $3,000	TOTAL WITH FORMAL PLANS
Manufacturing	22 8	31 1	13.7	37 6
Nonmanufacturing	43.8	64 0	28 2	67.5
All industries	25.7	35.6	15.7	41.7

only a small proportion of working income. By 1946 firms were 50% more likely to offer pensions to salaried workers than to wage earners. Nearly a third of nonmanufacturing firms, where lower-level workers were frequently salaried, had installed special pension plans for employees whose earnings exceeded the $3,000 Social Security ceiling—this was also true for 14% of manufacturing firms.

Finally, because federal policy did not require firms to bargain with unions over fringe benefits, employers tried to use benefits to discourage unionism. Industrial relations professionals were promoting private insurance as an antiunion device with the same rhetoric that had been used to sell welfare work. However, after unions won the right to bargain for benefits in the late forties, employers switched to support for Social Security increases, and they helped to win expanded Social Security benefits in the early fifties (Stevens 1990).

Public Policy and Union Support for Fringe Benefits

A number of analysts have argued that union agitation for coverage accounts for the remarkable increases in health and pension insurance during the war. Unfortunately, the NICB's studies do not present detailed data on the coverage of union workers, although there is some evidence that unionization was related to the presence of pension and health insurance (see tables 1 and 3). First, if we look cross-sectionally at 1935 we see that nonmanufacturing industries were more highly unionized (25% of all firms were unionized) than were manufacturing industries (10% of all firms were unionized), and they were more likely to offer pension insurance (31% vs. 7%). Figures for health insurance were similar across the two sectors, despite the fact that many nonmanufacturing firms still depended on mutual benefit associations or salary maintenance. Second, increases over time in unionization are related to the number of

TABLE 4

PERCENTAGE OF FIRMS REPORTING UNION CONTRACTS, BY SECTOR AND SIZE

	1928	1935	1939	1946
Manufacturing 4 7		9 6	43.0	80 2
Nonmanufacturing 3 8		25 0	45.3	59.5
Small firms 4.7		7.3	32 3	67.8
Large firms 5.5		13.9	48.6	80.5
All industries/firms .. . 4 9		11.7	43.4	77.3

NOTE —The 1928 sectoral figures are for small firms and the 1928 figure for all industries contains a higher proportion of small firms Small firms are those with fewer than 250 employees, large firms have 250 or more employees

employers who offer insurance. Between 1935 and 1946 unionization rose from 10% to 80% among surveyed manufacturing firms, and it rose from 25% to only 60% in nonmanufacturing (see table 4). Over that period both types of insurance underwent huge gains in manufacturing, but only pension insurance increased in nonmanufacturing areas (see table 1).

Unionization increased substantially in the aggregate over these years, and unions fought hard to win employment-related insurance even as they fought for the expansion of Social Security. Contemporary events suggest that public policy channeled union energies in the direction of private coverage. First, case law surrounding the Wagner Act entangled unions in a battle over the right to bargain for insurance coverage. The Wagner Act required firms to bargain with unions on the "wages and conditions of employment" without specifying whether "conditions" included benefits. Employers contended that "conditions" did not include benefits, and the courts backed them up until after the war, spurring a battle in the courtroom and at the workplace (Munts 1967, p. 10; Bernstein 1972). Unions, particularly the new Congress of Industrial Organizations (CIO) unions in the mass-production sector, fought the fringe-benefit battle in the courts by demanding the legal right to represent members in negotiations for benefits and, at the organizational level, by demanding that employers include benefits in union contracts (Bernstein 1972; Quadagno 1988). Labor historians (e.g., Bernstein 1972) seem to concur that if the Wagner Act had not made benefits a point of contention, unions might have spent more of their energies backing the expansion of Social Security.

The sharp rise in insurance between 1939 and 1946, therefore, cannot be attributed to union-negotiated benefit plans.[12] Only 4% of large firms

---

[12] By 1947 an estimated 1.5 million Americans were covered by union-negotiated health plans, some 90% of which were in the mining industry (which had experienced

in the 1946 survey had union contracts which included any sort of insurance scheme, while 81% were unionized and 67% offered health insurance (NICB 1947). However there is strong evidence that firms frequently adopted fringe benefits either to subvert union organizing efforts or to appease unions, and, in fact, a number of unions won contractual promises that employers would not discontinue their unilateral benefit plans (Bernstein 1972; Munts 1967, p. 49).

Second, the form group pension insurance took after passage of Social Security legislation also contributed to union lobbying for expanded employment-related coverage. The 1939 amendments made it clear that even if some Democrats envisioned Social Security as a full-coverage pension program, in practice it served as a safety net that was inadequate without supplementary coverage. By 1939 most firms had adopted some sort of graduated private pension program that would make up the difference between Social Security benefits and a certain percentage of preretirement income (NICB 1939).

These programs were marketed by insurance companies, who devised "integrated" pension plans that guaranteed a certain combined public-private retirement wage in the hope that promised increases in public benefit levels would decrease projected insurance industry expenditures (Achenbaum 1986, p. 46). The structure of Social Security legislation and the projections for future benefit increases thus shaped the insurance industry strategy. The nature of these plans in turn influenced union strategy, because they had the effect of nullifying Social Security increases. For each dollar increase in monthly Social Security benefits, private supplemental benefits would decrease by a dollar. The 1942 Revenue Act then sanctioned this form of supplemental coverage by making it tax exempt, while it rendered other discriminatory pension programs taxable. These formulas encouraged the rank and file to fight for increases in the total benefit levels guaranteed by their private pension programs because they would gain nothing from Social Security increases (see Quadagno 1988). Thus the structure of Social Security benefits led unions to promote private fringe-benefit programs at the organizational level.

Union support for private pension insurance had in part been stimulated by congressional resistance to the idea of expanding Social Security. Likewise union support for employment-related health insurance was reinforced by the failure of a series of public health insurance bills in Congress. Early proponents of Social Security wanted health coverage

---

strikes and a federal takeover during the war) and the textile/clothing industry (where industrywide jointly managed funds were established in 1942 contracts, Goldmann 1948, p. 47; Munts 1967, pp. 13–28) Those successes are evident in table 1.

included in the 1935 legislation, but acquiesced to fiscal conservatives in Congress. In 1939 Senator Robert Wagner introduced a bill that would have provided governmental health coverage for the needy, but that bill never came to a vote. In 1943 the Wagner-Murray-Dingell bill proposed universal and comprehensive health coverage to be administered under Social Security, but strong opposition from the American Medical Association (AMA) doomed it. In 1944 Roosevelt promoted an "economic bill of rights," which included the right to adequate medical care, and shortly after the end of the war Truman asked for national health insurance to cover all Americans, but the AMA defeated that bill as well (Anderson 1985; Berkowitz and McQuaid 1980; Starr 1982). Each defeat bolstered union support for private health coverage.

In sum, unions might have been expected to throw all of their support behind social insurance in the forties, but the Wagner Act and congressional opposition to the growth of Social Security kept them fighting for fringe benefits as well. When unions won the right to bargain over benefits in the 1947 Inland Steel decision, they did not give up on fringe benefits; on the contrary, having just lost ground in the Taft-Hartley Act, unions were anxious to demonstrate their newfound powers by winning fringe-benefit increases.

Alternative union strategies.—Historical factors produced unusual union strategies in certain sectors. These strategies led to substantially different kinds of benefit systems. An examination of several examples suggests that had public policy created a different set of incentives for unions in all industries, the outcome might have been a very different system of pension and health coverage throughout the economy.

First, AFL and CIO unions pursued markedly different strategies when it came to health benefits. The skilled AFL unions had traditionally sought to maximize their control over such matters as hiring and apprenticeships, and many AFL leaders still preferred to control member benefits directly through their own pension schemes and mutual benefit associations (Brandes 1976; Quadagno 1988, p. 160). By contrast the industrial CIO unions generally sought to formalize employment rights, and they backed employer-provided insurance schemes with nonrevokable rights to benefits. Consequently, mutual benefit associations survived in AFL-dominated industries, and health insurance prevailed in CIO-dominated industries. Those relationships became stronger during the war, despite the overall decline in welfare programs and the rise of insurance. Note that AFL unionism correlates .24 with mutual benefit association usage in 1935 (before the establishment of the CIO), .33 ($P < .10$) in 1939, and .40 ($P < .05$) in 1946 (see table 5). By contrast, CIO unionism correlates .29 with health insurance coverage in 1939, and .61 ($P < .01$) in 1946,

TABLE 5

CORRELATIONS OF AFL AND CIO UNIONISM WITH HEALTH COVERAGE

	1935	1939		1946	
	Unionism	AFL	CIO	AFL	CIO
Mutual benefit association .. .	24	33†	−.03	39*	15
Health insurance	12	.14	.29	12	.61**

NOTE —The CIO did not exist in 1935
† P < 10
* P < 05.
** P < 01

at which time it also correlates .38 ($P < .05$) with hospitalization insurance.[13] Here the different strategies of AFL and CIO unions had clear effects on fringe-benefit development, even if unions did not yet receive contractual rights to negotiate for health insurance.

Second, the railway industry shows evidence of what American pension and health coverage might have looked like had unions in more industries pursued the strategy of providing insurance themselves. Evidence from railroads and mining suggests that union-sponsored benefits precluded the adoption of employer-provided insurance. Railway unions had first been organized as benevolent societies and had long provided death, sickness, and old-age benefits for members. By 1931, railway unions paid out more in sickness benefits and in old-age pensions than did unions in any other industry—half as much as all other unions combined (Bureau of Labor Statistics 1933, pp. 313–31). Union health coverage provided an alternative to company benefits, hence table 1 shows that among transportation/communication firms, company pensions and health insurance *declined* between 1935 and 1939. By 1946 transportation reported the lowest incidence of offering health insurance of any industry, with the exception of banking, where salary maintenance programs were still used as a substitute. In contrast, mining unions had provided death benefits, but seldom offered pension or sickness coverage, and table 1 indicates that pension and health insurance had expanded significantly for miners by 1946. The success of union-provided insurance among railways suggests that if unions in other sectors had installed

[13] Correlations use the original industrial categories from the NICB reports. Note that only industry-level data were available, hence these correlations represent the relationships between unionism and benefits among industries, not firms In order to maximize the number of cases in each panel (26 in 1935, 25 in 1939, and 32 in 1946) I have not regrouped the industries to make them comparable over time.

insurance schemes before corporations did so, as Samuel Gompers advocated, union-provided insurance schemes might have proliferated during the war and become the norm in American industry.

CONCLUSION

Industrial Arguments

Data collected by the NICB largely contradict the received wisdom that industrial factors and wartime public policies stimulated the growth of fringe-benefit programs. Early interindustry differences suggested that firms with skilled workers were most likely to adopt fringe-benefit packages and that organizational scale was associated with an industry's offering fringe benefits. Yet the experiences of the thirties tend to discount industrial arguments, because despite the fact that the problems normally associated with unionism—labor turnover, organizational scale, and labor-market segmentation—were mitigated during this period of economic crisis, the incidence of pension and health insurance increased. Furthermore, interindustry data from the forties suggest that these industrial problems were not related to increases in the prevalence of fringe benefits.

Institutional Theory and Interest Group Goals

I have developed an alternative institutional explanation that links public policy to interest group goals and in turn to organizational outcomes. The findings presented here underscore the power of an institutional approach that treats interest group goals as an intermediate variable between environmental context and organizational outcomes. As expected, union and business groups did not pursue consistent, predictable goals in the realms of health and pension coverage; rather, public policy shifts caused union and business goals to vary over time, and those groups were clearly instrumental in popularizing private insurance. Previous organizational studies have provided important evidence of the effects of the policy environment on group goals and organizational outcomes (Fligstein 1990; Baron et al. 1986), and here I have tried to extend that perspective by looking at the effects of public policy on the shifting goals of all major groups. In addition to providing a contextual framework for understanding group interests, this approach highlights the importance of recognizing that a group may pursue a single end alternately at the organizational level and at the political level.

The broad institutional context of American industrial and state development clearly contributed to union and business preferences for private

forms of coverage. America's early industrial development led to the rise of sophisticated corporate bureaucracies; however, the development of the federal bureaucracy was comparatively slow (Chandler 1990; Skowronek 1982). As a result firms and unions experimented with private insurance long before Washington considered getting into the act, and when these groups discovered labor control functions in private coverage they became reluctant to relinquish control to the state. In Philip Selznick's (1948) terms, these practices had developed organizational constituencies that would fight for them in the future. This set the groundwork for the situation in the teens and twenties, when both business and labor leaders advocated private forms of insurance. Weak public protections had also stimulated the development of a large and sophisticated insurance industry that would promote employment-related health and pension coverage during the economic downturn of the thirties.

The strength and consistency of union support for private insurance is particularly surprising. Before the early thirties, the AFL leadership supported union-provided coverage instead of public coverage, because they believed that coverage increased member loyalty. After 1935 unions fought for employer-provided insurance because public policy made union bargaining rights a point of contention. Finally, during the forties the popularity of federally approved supplemental pension insurance plans bolstered union support for the expansion of private benefits because the supplemental plans canceled out increases in public coverage.

Intended and Unintended Consequences of Public Policy

Public policies that were expected to increase the incidence of private insurance did not always succeed. There were, however, some policies that did achieve this end without being designed to do so. Wartime tax and wage policies were certainly intended to help diffuse employer-provided insurance, but by tracing the empirical effects of those policies I have shown that they appeared after the wartime rise in fringe benefits. Instead, Social Security legislation boosted employment-related pension insurance by institutionalizing retirement without providing an adequate retirement wage. The original Social Security Act of 1935 had a marked effect on the prevalence of private pensions in high-wage sectors and also caused firms to rewrite their pension plans. Moreover, the 1939 amendments to the act gave a substantial boost to private insurance schemes because they expanded eligibility without raising contributions. Social Security revenues could not support the benefit increases that Roosevelt had originally hoped for, and when benefits were decimated by inflation, corporations responded by installing private supplementary pension programs. The failure of several proposals to extend Social Secu-

rity to include health insurance had a similar effect on private health/ accident coverage.

The demise of welfare programs was in part wrought by these same policies, not by the economic exigencies of the Depression era. Health and pension forms of welfare work rose during the early Depression. But these forms declined after 1935 as firms responded to Social Security legislation by replacing their antiquated informal pension programs with modern supplementary insurance. Both forms of welfare coverage fell prey to the struggle over the interpretation of the Wagner Act, which led firms to offer insured fringe benefits outside of the union bargaining process.

Finally, while this article has focused on the factors that contributed to the growth of private employment-related insurance, the perspective employed here can inform institutional studies of state policy formulation as well. It should be clear, for instance, that the factors that shaped labor, business, and insurance industry goals vis-à-vis firm-level insurance coverage also shaped their goals vis-à-vis social insurance legislation, and the events of the forties clearly set the course for the future of social insurance (Amenta and Skocpol 1988). While some institutional studies of politics have analyzed how institutional factors affect the capacities of different groups to prevail in political struggles (Weir and Skocpol 1983), institutionalists have largely ignored how institutional context shapes the goals groups pursue in the first place.

REFERENCES

Achenbaum, W. Andrew. 1986. *Social Security. Visions and Revisions*. Cambridge: Cambridge University Press.
Amenta, Edwin, and Theda Skocpol. 1988. "Redefining the New Deal: World War II and the Development of Social Provision in the United States." Pp. 81–122 in *The Politics of Social Policy in the United States,* edited by Margaret Weir, Ann Shola Orloff, and Theda Skocpol. Princeton, N.J.· Princeton University Press.
Anderson, Odin W. 1985. *Health Services in the United States: A Growth Enterprise since 1875*. Ann Arbor, Mich.: Health Administration Press.
Baker, Helen. 1940. *Current Policies in Personnel Relations in Banks*. Princeton, N.J.: Princeton University, Industrial Relations Section.
Baron, James N , Frank R Dobbin, and P. Devereaux Jennings. 1986. "War and Peace. The Evolution of Modern Personnel Administration in U.S. Industry." *American Journal of Sociology* 92:350–83.
Baron, James N., P. Devereaux Jennings, and Frank R. Dobbin. 1988. "Mission Control The Development of Personnel Systems in U.S. Industry." *American Sociological Review* 53:497–514.
Berkowitz, Edward, and Kim McQuaid. 1980. *Creating the Welfare State: The Political Economy of Twentieth-Century Reform* New York Praeger.
Bernstein, Irving. 1972. *The Turbulent Years: A History of the American Worker, 1933–1941* Boston Houghton-Mifflin.

Blau, Peter M., and Richard A. Schoenherr 1971. *The Structure of Organizations.* New York Basic.

Brandes, Stuart D. 1976. *American Welfare Capitalism, 1880–1940.* Chicago: University of Chicago Press.

Brody, David. 1980. *Workers in Industrial America.* New York: Oxford University Press.

Bureau of Economic Analysis. 1986. *The National Income and Product Accounts of the United States, 1929–1982.* Washington, D.C.: Government Printing Office.

Bureau of Labor Statistics. 1928. "National Conference on Mutual Benefit Associations." *Monthly Labor Review* 27:104–5.

———. 1933. "Insurance, Benefit, and Pension Plans." *Monthly Labor Review* 36:311–15.

———. 1932–43. *Monthly Labor Review.*

Bureau of the Census. 1940. *Sixteenth Census of the Population,* vol. 1. Washington, D.C.: Government Printing Office.

———. 1947. *Census of Manufactures, 1947,* vol. 6. Washington, D.C.. Government Printing Office.

Carter, Michael J., and Susan B. Carter. 1985 "Internal Labor Markets in Retailing: The Early Years." *Industrial and Labor Relations Review* 38:586–98.

Chandler, Alfred D., Jr. 1990. *Scale and Scope. The Dynamics of Industrial Enterprise.* Cambridge, Mass.: Harvard University Press.

Civilian Production Administration. 1947. *Industrial Mobilization for War.* Washington, D.C.: Government Printing Office.

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48:147–60.

———. 1991. Introduction to *The New Institutionalism in Organizational Analysis,* edited by W. Powell and P. DiMaggio. Chicago· University of Chicago Press.

Doeringer, Peter B., and Michael J. Piore. 1971. *Internal Labor Markets and Manpower Analysis.* Lexington, Mass.: Heath.

Edwards, Richard. 1979. *Contested Terrain.* New York: Basic.

Esping-Andersen, Gosta. 1990. *The Three Worlds of Welfare Capitalism.* London: Polity.

Fligstein, Neil. 1990. *The Transformation of Corporate Control.* Cambridge, Mass.: Harvard University Press.

Goldmann, Franz. 1948 *Voluntary Medical Care Insurance in the United States.* New York: Columbia University Press.

Gordon, David M., Richard Edwards, and Michael Reich. 1982. *Segmented Work, Divided Workers.* London: Cambridge University Press.

Hall, Peter. 1986. *Governing the Economy: The Politics of State Intervention in Britain and France.* New York Oxford University Press.

Hodson, Randy. 1978. "Labor in the Monopoly, Competitive, and State Sectors of Production." *Politics and Society* 8:429–80.

Ilse, Louise. 1953. *Group Insurance and Employee Retirement Plans.* New York: Prentice-Hall.

Insurance Almanac. 1930. "Health Insurance in 1929 " P. 1189 in *Insurance Almanac.* Englewood, N.J.: Underwriter.

———. 1932. "Health Insurance in 1931." P. 1016 in *Insurance Almanac.* Englewood, N.J.: Underwriter.

Jacoby, Sanford. 1984. "The Development of Internal Labor Markets in American Manufacturing Firms." Pp. 23–69 in *Internal Labor Markets,* edited by Paul Osterman. Cambridge, Mass.: MIT Press.

———. 1985. *Employing Bureaucracy: Managers, Unions, and the Transformation of Work in American Industry, 1900–1945.* New York: Columbia University Press.

James, Marquis. 1947. *The Metropolitan Life· A Study in Business Growth*. New York· Viking.

Kerry, Tom. 1980. *Workers, Bosses; and Bureaucrats: A Socialist View of Labor Struggles since the 1930s*. New York: Pathfinder.

Krasner, Stephen D. 1984. "Approaches to the State: Alternative Conceptions and Historical Dynamics." *Comparative Politics* 17.223–46

Kuznets, Simon. 1961. *Capital in the American Economy*. National Bureau of Economic Research. Princeton, N.J. Princeton University Press.

Labor Statistics Bureau (Bureau of Labor Statistics) 1917. *Proceedings of the Employment Managers' Conference, Philadelphia, April 2 and 3, 1917*. Labor Statistics Bureau Bulletin 227. Washington, D.C.: Government Printing Office.

Landers, Jackson K. 1935. "How Accident Insurance Protects Incomes." *Rough Notes* 78:13–14.

Macaulay, Hugh. 1959. *Fringe Benefits and Their Federal Tax Treatment* New York: Columbia University Press.

Meyer, John W., and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." *American Journal of Sociology* 83.340–63.

Munts, Raymond 1967. *Bargaining for Health. Labor Unions, Health Insurance, and Medical Care*. Madison. University of Wisconsin Press.

National Industrial Conference Board (NICB). 1925. *Industrial Pensions in the United States*. New York: NICB

———— 1929. *Industrial Relations in Small Plants*. New York. NICB.

———— 1934. *Recent Development in Industrial Group Insurance*. New York. NICB.

————. 1936. *What Employers Are Doing for Employees*. New York: NICB.

————. 1937. *Personnel Practices Governing Factory and Office Administration*. New York: NICB.

————. 1939. *Company Pension Plans and the Social Security Act*. New York: NICB.

————. 1940a. *Personnel Activities in American Business*. New York: NICB.

————. 1940b. *Company Pension Plans and the Amended Social Security Act— Complete Texts and Digests*. New York NICB

————. 1947. *Personnel Activities in American Business, rev. ed*. New York. NICB.

———— 1954. *Personnel Practices in Factory and Office*. New York: NICB.

National Personnel Association (NPA). 1922. *Pensions for Industrial and Commercial Employees* New York· NPA.

Nelson, Daniel. 1975. *Managers and Workers*. Madison University of Wisconsin Press.

Parks, Donald S. 1936. "1936 Personnel Trends." *Factory Management and Maintenance* 94.39–40.

Pelling, Henry. 1960. *American Labor*. Chicago. University of Chicago Press.

Perrow, Charles. 1986. *Complex Organizations: A Critical Essay*, 3d ed. New York: Random House.

Pugh, D S , D. J. Hickson, C. R. Hinings, and C Turner. 1969. "Dimensions of Organizational Structure." *Administrative Science Quarterly* 14:91–114.

Quadagno, Jill. 1984. "Welfare Capitalism and the Social Security Act of 1935." *American Sociological Review* 49:632–47.

———— 1988. *The Transformation of Old Age Security. Class in Politics in the American Welfare State*. Chicago: University of Chicago Press.

Rein, Martin. 1982. "The Social Policy of the Firm." *Policy Sciences* 14:117–35.

Romasco, Albert U 1983. *The Politics of Recovery: Roosevelt's New Deal*. New York. Oxford University Press.

Rough Notes. 1935. "Sell Accident or Disability Insurance: Dig a Gold-Nugget from Your File of Present Clients." *Rough Notes* 78 11–23.

Schieber, Sylvester J. 1982. *Social Security*. Washington, D.C.: Employee Benefit Research Institute.

Schlesinger, Arthur M., Jr. 1958. *The Coming of the New Deal.* Boston. Houghton Mifflin.

Seidman, Joel 1953. *American Labor from Defense to Reconstruction.* Chicago· University of Chicago Press.

Selznick, Philip. 1948. "Foundations of the Theory of Organization." *American Sociological Review* 13:25–35

Skowronek, Stephen 1982. *Building a New American State. The Expansion of National Administrative Capacities, 1877–1920.* Cambridge: Cambridge University Press.

Slichter, Sumner. 1961. *Potentials of the American Economy: Selected Essays of Sumner H. Slichter.* Cambridge, Mass.. Harvard University Press.

Starr, Paul. 1982 *The Transformation of American Medicine.* New York: Basic.

Stevens, Beth. 1988. "Blurring the Boundaries: How the Federal Government Has Influenced Welfare Benefits in the Private Sector." Pp. 123–48 in *The Politics of Social Policy in the United States,* edited by Margaret Weir, Ann Shola Orloff, and Theda Skocpol. Princeton, N.J.. Princeton University Press.

———. 1990. "Labor Unions, Employee Benefits, and the Privatization of the American Welfare State." *Journal of Policy History* 2:233–60.

Stone, Mildred. 1957. *Since 1845: A History of the Mutual Benefit Life Insurance Company.* New Brunswick, N.J.: Rutgers University Press.

Tiger, E. S. 1932. "'Income Insurance' Cornerstone." *Rough Notes* 75:11–13.

Tolbert, Pamela, and Lynne G. Zucker. 1983. "Institutional Sources of Change in the Formal Structure of Organizations· The Diffusion of Civil Service Reforms, 1880–1935." *Administrative Science Quarterly* 23.22–39.

Weir, Margaret, and Theda Skocpol. 1983. "State Structures and Social Keynesianism." *International Journal of Comparative Sociology* 24.4–29.

Witte, Edwin. 1962. *The Development of the Social Security Act.* Madison. University of Wisconsin Press

Wolman, Leo. 1936. *Ebb and Flow in Trade Unionism.* New York: National Bureau of Economic Research.

Zelizer, Viviana. 1979. *Morals and Markets: The Development of Life Insurance in the United States.* New York. Columbia University Press.

Zysman, John. 1983. *Governments, Markets, and Growth: Financial Systems and the Politics of Industrial Change.* Ithaca, N.Y.: Cornell University Press.

Commentary and Debate

To conserve space for the publication of original contributions to scholarship, the comments in this section must be limited to brief critiques. They are expected to address specific errors or flaws in articles and reviews published in the *AJS*. Comments on articles are not to exceed 1,500 words, those on reviews 750 words. Longer or less narrowly focused critiques should be submitted as articles. Authors of articles and reviews are invited to reply to comments, keeping their replies to the length of the specific comment. The *AJS* does not publish commenters' rebuttals to authors' replies. We reserve the right to reject inappropriate or excessively minor comments.

PACS, INTERLOCKS, AND REGIONAL DIFFERENCES IN CORPORATE CONSERVATISM: COMMENT ON CLAWSON AND NEUSTADTL

In a recent article (*AJS* 94 [January 1989]: 749–73), Dan Clawson and Alan Neustadtl examined variation in the conservatism of corporate PACs in the 1980 election. Their research replicated many of the findings of my study of the 1982 election (Burris 1987, pp. 732–44) but differed on two key points. Clawson and Neustadtl found greater conservatism among corporations with the fewest director interlocks. They also found no significant regional differences in corporate political behavior, whereas I found Sunbelt (South and West) firms to be more conservative than Frostbelt (Northeast and Midwest) firms. Clawson and Neustadtl acknowledged that these results might reflect differences in the two elections studied, but argued that methodological "limitations" of my study led them to question my findings for the 1982 election. Below I refute Clawson and Neustadtl's criticisms and further argue that (1) the behavior of corporate PACs in 1982 was indeed different than in 1980; (2) the evidence for a causal association between interlocks and political behavior is weaker than Clawson and Neustadtl suggest; and (3) regional differences in corporate political behavior were significant in both 1980 and 1982.

Clawson and Neustadtl's first criticism of my study is that it is based on a limited sample of corporations. This is an odd claim since my sample included 443 corporations, while theirs included only 243. Clawson and Neustadtl asserted that my sample excluded large, private nonindustrial firms such as Cargill and Bechtel. This statement is incorrect. My sample did include Cargill and Bechtel and every other nonindustrial corporation listed by *Ward's Directory* as among the 100 largest private companies. My 443 firms represent all of the largest 1,000 U.S. corporations with PAC contributions of at least $5,000. Clawson and Neustadtl's 243 firms represent all U.S. corporations, regardless of size, with contributions of at least $25,000. Of the 243 firms in Clawson and Neustadtl's sample, only 19 were excluded from my sample because of their small size. By comparison, Clawson and Neustadtl's more restrictive criterion of $25,000 in PAC contributions resulted in the exclusion from their sample of approximately 200 firms—including such giants as Goodyear, Phillips Petroleum, Texas Instruments, Reynolds Metals, Ralston-Purina, PPG Industries, and American Motors.

More important than sample size per se is the question of whether or not a sample is representative of corporations as a whole. Comparing the corporations in my 1982 sample with $25,000 or more in PAC contributions with those contributing less than $25,000, I found that the former group was significantly less conservative than the latter and significantly more likely to contribute to incumbents. In other words, the big-budget PACs that are the exclusive focus of Clawson and Neustadtl's study exhibit different political behavior than the smaller-budget PACs that are excluded from their sample. Thus, their sample is not just small, it violates the first principle of sampling: it selects on the basis of a variable that is correlated with the dependent variable.

Clawson and Neustadtl's second criticism of my study concerns the sample of corporations that I use as the reference group for counting director interlocks. My measure of interlocks, based on a 1980 study, counts the number of interlocks with a sample of 100 large firms. Their measure, based on a 1976 study, counts the number of interlocks with the 200 largest nonfinancial and 50 largest financial firms. Their sample has a slight advantage in size; mine has a slight advantage in being more up-to-date. The chief difference is that their sample is chosen purely on the basis of company size, whereas mine is a stratified sample that includes the dominant firms in each of 18 industry groups. Clawson and Neustadtl object to this method because it omits some very large firms and includes slightly smaller firms from industries that would otherwise be underrepresented in the sample. In fact, this is precisely why I chose the 1980 sample over the 1976 sample. According to Useem (1984)—the theorist we both cite as authority for the claim that director interlocks

influence corporate politics—interlocks enhance classwide rationality insofar as they involve directors in the affairs of firms in *diverse* sectors of the economy. The more equal representation of industries in the 1980 sample makes it well suited for testing this hypothesis. The consequences of choosing one measure over the other should not be exaggerated. Using data on the 1976 sample provided by Beth Mintz, I found a correlation of 0.84 between the number of interlocks in the two samples. This is not much different from Clawson and Neustadtl's own estimate of a correlation of 0.9 between two identically constructed samples taken four years apart.

Clawson and Neustadtl's third criticism of my study concerns my measure of corporate support for New Right candidates. To measure corporate conservatism, I identified a group of candidates supported by at least three of the five biggest New Right PACs; I then computed the share of each corporation's contributions that went to these right-wing candidates. Clawson and Neustadtl note that these New Right candidates are disproportionately from the Sunbelt. There is a modest tendency for corporations to contribute more heavily to candidates in their own region. From this Clawson and Neustadtl conclude that my measure is "biased" in a manner that predisposes me to find that Sunbelt firms contribute more heavily to conservative candidates. This charge of measurement bias is misplaced. The fact that Sunbelt corporations are located in geographic proximity to the districts contested by New Right candidates may be one *reason* they support those candidates, but it does not indicate a bias in my measure. Sunbelt corporations may support conservative candidates because they are local, or they may support local candidates because they are conservative. In either case, the candidates in question are still the ones most favored by conservative strategists.

Clawson and Neustadtl compute their conservatism measure differently. They assign each contribution a score of +1 if it was to a candidate backed by a New Right PAC and −1 if it was to an opponent of a New Right candidate. Aggregating these scores across 13 New Right PACs and averaging them over the contributions of each firm yields a scale that ranges from a potential high of +13 (most conservative) to a potential low of −13 (most liberal). Explaining their rationale, they say in their article: "We . . . took behavior similar to conservative ideological PACs as evidence of conservatism and behavior opposed to them as evidence of liberalism" (p. 753). Contrary to what Clawson and Neustadtl say, this measure does not eliminate the regional "bias" that they ascribe to my measure. In the abstract, each New Right candidate in a region is matched by an opponent to whom corporations may also contribute, so that a propensity to support local candidates has an equal chance of raising or lowering a firm's score on their scale. In reality, however, the

chances are far from equal. The types of candidates favored by New Right PACs have a much higher probability of receiving corporate support than the types of candidates opposed by New Right PACs. This is evidenced by the fact that the average score of firms on Clawson and Neustadtl's conservatism/liberalism scale is significantly higher than zero. If Clawson and Neustadtl believe that the uneven geographic distribution of New Right candidates is a problem for my measure of conservatism, they should recognize that their measure is subject to the same problem.

When I constructed my conservatism measure, I debated whether I should focus exclusively on contributions to New Right candidates or also count (as evidence of liberalism) contributions to the opponents of New Right candidates. The reason I decided against the latter method is that, contrary to Clawson and Neustadtl, I believe it is mistaken to treat contributions to New Right candidates and contributions to their opponents as symmetrical acts of conservatism and liberalism. Contributing to the opponent of an ultraconservative is not the same thing as contributing to a liberal candidate or to a candidate favored by liberal strategists. Clawson and Neustadtl say that identifying candidates who are opposed by New Right PACs "enables us to differentiate an extreme liberal . . . from a middle-of-the-road moderate" (Clawson, Neustadtl, and Bearden 1986, p. 800). In fact, most candidates opposed by New Right PACs are middle-of-the-road Democrats. These are not usually the most liberal Democrats, since liberal Democrats typically seek office in liberal districts where the ultraconservative Republicans favored by the New Right have little chance of nomination. In 1982, for example, Democratic incumbents who faced opponents financed by the National Conservative PAC (NCPAC), the flagship of the conservative PACs, were slightly *less* liberal on average (as measured by *National Journal* ratings) than Democratic incumbents who did not face NCPAC-financed opponents. The liberal group, Americans for Democratic Action (ADA), endorsed approximately the same number of candidates in 1982 as those backed by NCPAC, but these were not typically the opponents of NCPAC candidates. Only about one-fourth of ADA candidates ran against NCPAC opponents and vice versa—further evidence that supporting liberals is not the same thing as opposing ultraconservatives. More to the point, most corporate contributions to opponents of New Right candidates are motivated not by liberalism but by *pragmatism*. Corporations give to Democratic incumbents because of their seniority, committee assignments, or probable reelection. The overall share of contributions to Democratic incumbents is a useful index of political moderation; however, there is little reason to treat contributions to the opponents of New Right candidates as more "liberal" in motivation than contribu-

tions to other Democrats, since the criteria by which corporations decide which Democrats to support are largely independent of political ideology.

Next, let me turn to the evidence that Clawson and Neustadtl present for their claim that director interlocks influence corporate political behavior. On one point we agree: at the zero-order level there is a modest negative correlation between interlocks and political conservatism. At issue is whether this demonstrates the causal importance of interlocks or of some other variable correlated with interlocks. Clawson and Neustadtl admit that there is a high correlation (approximately 0.6) between director interlocks and company size as measured by sales, but they maintain that the number of director interlocks remains a significant predictor of political behavior after effects of size are controlled for. This claim is suspect for two reasons. First, Clawson and Neustadtl's limited sample is not representative of the overall distribution of firms by size, interlocks, and political behavior. Second, Clawson and Neustadtl assume that the relationship between size and political behavior must be linear. For a variable such as sales, which has a very skewed distribution, a logarithmic transformation should always be considered. Had Clawson and Neustadtl measured size by the logarithm of sales, they would have found a much stronger correlation between size and political behavior in 1980. Using a larger sample and measuring size by the logarithm of sales, I found that interlocks did not have a significant effect on political behavior in 1980 after size was controlled for (Burris and Salt 1990).

Finally, let me comment on the question of regional differences in corporate political behavior. This issue has been thoroughly examined in a recent paper by Salt (1989), so I will limit myself to a summary of several key findings. In my study of the 1982 election, I concluded that the simple Sunbelt/Frostbelt dichotomy gives an oversimplified view of regional differences in corporate politics and recommended that future research examine more narrowly defined regions. Had Clawson and Neustadtl done so, they would have found highly significant regional differences in the 1980 election. As Salt shows, Sunbelt firms were more conservative than northeastern firms (and more willing to support nonincumbents) in every election from 1978 to 1986. The problem with the simple Sunbelt/Frostbelt dichotomy is that midwestern firms were also very conservative, especially in 1978 and 1980. For these years, collapsing the conservative Midwest with the liberal Northeast produces a heterogeneous category that, on average, was neither more nor less conservative than the Sunbelt. In 1982, midwestern corporations moved dramatically toward a more moderate political strategy. Their average contributions to New Right candidates declined more sharply than those of any other region, as did their contributions to (mostly Republican) nonincumbents. This may reflect some disaffection with right-wing poli-

cies, although my guess is that it was mostly a tactical response to the severity of the 1982 recession in the Midwest and the consequent vulnerability of Republican candidates in that region. As a result of this shift, the Sunbelt/Frostbelt dichotomy became a significant predictor of political behavior for the first time in 1982. This association holds for numerous measures of political behavior, including several (such as the percentage of contributions to incumbents) over which there can be no disagreement as to how they should be computed.

This points to the differences between 1980 and 1982. By historical standards, the 1980 election was characterized by an unusually high level of corporate support to conservative candidates—especially challengers. On average, northeastern firms and larger (more highly interlocked) firms were a bit more cautious as to how fully they supported this conservative shift. In 1982, a by-election year in the midst of a recession, business assumed a more defensive political posture and the range of corporate political behavior narrowed and shifted toward the center. The modest differences between large and small firms decreased, while new regional alignments emerged. Clawson and Neustadtl's 1980 findings differ from my 1982 findings in part for the simple reason that business political behavior is responsive to changing historical circumstances.

VAL BURRIS

University of Oregon

REFERENCES

Burris, Val. 1987. "The Political Partisanship of American Business." *American Sociological Review* 52:732–44.
Burris, Val, and James Salt. 1990. "The Politics of Capitalist Class Segments." *Social Problems* 37:341–59.
Clawson, Dan, Alan Neustadtl, and James Bearden 1986. "The Logic of Business Unity." *American Sociological Review* 51:797–811.
Salt, James. 1989. "Sunbelt Capital and Conservative Political Realignment in the 1970s and 1980s." *Critical Sociology* 16 (2–3): 145–63
Useem, Michael. 1984. *The Inner Circle* New York: Oxford University Press.

MYOPIC METHODS. MEASURES OF CONSERVATISM IN CORPORATE PAC ANALYSIS

The substantive issues at stake in this exchange are not enormous. One page of our 25-page article indicated how and why we differ from Burris; his multipage comment focuses almost exclusively on that one page. In our article we wrote that some of Burris's procedures *might* have been responsible for his finding that Sunbelt and Frostbelt corporations had

differing contribution patterns and implicitly suggested that Burris should check for this possible distortion in any future work. Burris's comment has considerable detail, but he has not addressed the issues we raised. His analysis differs from ours in the sampling criteria (economic vs. political), the year analyzed (1982 vs. 1980), and the measure of conservatism employed (New Right only vs. adjusting for contributions to their opponents). For this comment we have duplicated Burris's sample and rerun his analyses for his chosen year, 1982. We find that, when the controls we suggested are introduced, Burris's principal finding is no longer significant. This reply begins with a discussion of sampling issues, briefly mentions the differences in the two elections, then examines the problems with Burris's measure of conservatism.

Sampling

Our article's first criticism of Burris concerned sampling; in this comment Burris replies with a countercriticism of our sample. Burris presents his 1987 article as a test of corporate liberal theory, beginning with the statement in the abstract that "the four disconfirmed theories are all variants of a perspective known as the theory of 'corporate liberalism'" (p. 732). There are many versions of corporate liberalism, but the basic argument is that while most businesses, especially smaller businesses, are conservative, the largest and most central corporations (or at least a substantial number of them) have supported liberal reforms at key junctures. Burris defines his sample to include only the *economically* largest firms; he thus makes an a priori restriction of the possible variation in the independent variables appropriate to a test of corporate liberal theory. We define our sample to include the largest *political* actors, those who contributed $25,000 or more. This turns out to include almost entirely economically large firms, but for us this is a substantive finding, not a defining criterion.[1]

These issues are lost in Burris's comment where he states: "Clawson and Neustadtl's first criticism of my study is that it is based on a *limited* sample of corporations. This is an odd claim since my sample included 443 corporations, while theirs included only 243" (our emphasis). In fact, we wrote that Burris used "1982 data and a slightly *different* sample" (p. 760; our emphasis); we are aware that 443 is more than 243. Samples are not to be judged by the absolute number of cases; Burris's list of

---

[1] A relatively small firm with high political mobilization could easily qualify for our sample. Even an academic department could raise enough to do so: if 25 faculty each contributed an average of $10 a week, in a two-year election cycle, this would total more than $25,000 and thus qualify for our sample.

corporations that he includes and we exclude is simply irrelevant. We apparently did err in saying Burris's sample omitted Cargill and Bechtel, but this was because Burris *did not specify* which list of large, privately owned corporations he used. Because his other corporations came from the *Fortune* lists we assumed that his list of privately held corporations did as well. The *Fortune* listing does not include Cargill and Bechtel; in this comment Burris reports for the first time that he used *Ward's* list, which does include them.

Burris never considers the theoretical basis for selecting a sample, but he does offer a critique "Comparing the corporations *in my 1982 sample* with $25,000 or more in PAC contributions with those contributing less than $25,000, I found that the former group was *significantly* less conservative than the latter and *significantly* more likely to contribute to incumbents" (all emphases ours). Three comments are in order here. First, Burris fails to provide any actual numbers to support his claims, making it difficult to assess his reported findings. Second, even for the claim he has made he is wrong. In order to test his claims we replicated Burris's sample for 1982, the year he studied, and examined the percentage given to Republicans by the PACs in each of the two samples.[2] *Within his sample,* the $25,000-and-over PACs are less conservative than the under-$25,000 PACs, but this difference is *not* significant. Table 1 presents the actual figures: $25,000 and over, 66.7% to Republicans (SE = 1.2); under $25,000, 68.3% to Republicans (SE = 1.7%), one-tailed t-test not significant at $P < .05$. Third, more important, Burris has made the wrong comparison. What is to be said about Burris using *his* sample to test the representativeness of *our* sample? In an exchange specifically focused on the relative merits of the two samples, it is remarkable that Burris, without qualification or explanation, simply used his as the standard. Presumably the appropriate standard is to compare both samples with the population of all corporate PACs. As table 1 indicates, the mean percentage to Republicans in 1982 for the population of all corporate PACs is 63.3%; the PACs selected by our sampling criterion gave 65.1% to Republicans; Burris's sample gave 67.4% to Republicans. Our sample is actually closer to the population mean than Burris's.[3]

[2] Some may consider the data used in these two studies as samples and interpret the following tests of significance as comparisons of sample parameter estimates to a fixed population Others may consider the data as actually constituting two populations and be concerned about the appropriateness of using inferential tests. However, these tests may also be viewed as a test of the causal processes that produced these two data sets (Blalock 1979, pp. 241–43). Note that our conclusions do not depend on whether or not these tests are employed

[3] We are trying to give Burris every possible edge in this comparison; of course the correct evaluation of our sample would use 1980 data (since that was the focus of our

TABLE 1

A COMPARISON OF ALTERNATIVE SAMPLES OF CORPORATE PACS·
PERCENTAGE OF MONEY TO REPUBLICANS IN 1982

	Mean	SD	Number	t-test
Comparison *within* Burris's sample:				
PACs of $24,999 or less	68.3	23.9	205	
PACs of $25,000 or more	66.7	18.4	254	0.78
Comparison of *population* with two samples·				
Population of all PACs over $1.00	63 3	27 5	1,315	
Burris sample	67 4	21.0	459	3 25
Clawson/Neustadtl sample	65.1	19.9	337	1.22

Year of Election

We did not find a Frostbelt-Sunbelt split in 1980; for 1982, Burris found
no Frostbelt-Sunbelt difference in preference for Republicans versus
Democrats but did report a difference in giving to New Right candidates.
In our original article we noted that the differences in the two elections
might produce differing results, that corporate PACs gave 47% of their
Senate contributions to Republican challengers in 1980 but only 17% to
such candidates in 1982. Therefore we said "Burris's results may be
correct for 1982 and ours for 1980" (p. 760). In order to focus on the
important methodological differences between us and Burris, for pur-
poses of this reply we have accepted Burris's chosen year and run all
analyses for 1982, eliminating all differences based on election year.

Measures of Conservatism

The differences between Burris's results and our own depend not only
on differing samples and the year analyzed, but also on differences in the
way we measure conservatism. Burris adapted a measure we devised
(Clawson, Kaufman, and Neustadtl 1983, 1985; Clawson, Neustadtl,
and Bearden 1986) but changed it (see Burris 1987) in a way that created
the potential for serious problems. Our measure increases a corporation's
political conservatism score if it contributes to New Right candidates and
decreases its score if it gives to the opponent of these candidates. Burris
omits the decrease for opposing New Right candidates. Imagine a
scenario where a particular group of corporations gave substantially
more to *both* sides of the races involving New Right candidates. Using

article). In 1980, the PACs in our sample gave 65.2% to Republicans, and all corporate
PACs gave 63.8%; ours gave 58.2% to incumbents, all gave 58.6%.

our measure these corporations would not be identified as particularly favorable to conservatism; using Burris's measure they would be.

What Burris has done is equivalent to a proponent of corporate liberalism looking at the absolute amount of money given to Democrats, and reporting that the biggest corporations gave Democrats more than smaller corporations. This is undoubtedly true, but the biggest corporations also gave more to Republicans. They gave more in total. Similarly, Burris has constructed a New Right measure that only considers giving to New Right candidates and makes no adjustment for giving to their opponents. If corporations tend to give within their region, and if Burris's New Right candidates are located in the Sunbelt, his measure will identify Sunbelt corporations as more conservative even if they give as much to New Right opponents as they do to New Right candidates.

This is not simply an abstract possibility. Corporations have a strong tendency to favor races in their geographic area. In a survey of corporate PAC directors we conducted in 1986 (response rate 58%), 60% of the responding corporations said that if a candidate represented a state or district where the company had a facility this would have "a great deal" of influence on whether or not they contributed, and a further 25% said that it would have "a moderate amount" of influence. A large majority of the "New Right" candidates Burris identified were in the Sunbelt. In our article we simply noted that "it is quite plausible that" this explains the modest difference Burris finds between Sunbelt and Frostbelt corporations (p. 761).

To examine the issue we raised we recreated Burris's data set and reran his analysis of contributions to New Right candidates by Sunbelt versus Frostbelt corporations using his definition of New Right candidates. We found that, just as Burris reports, Sunbelt corporations gave more of their money to New Right candidates (22.18%) than did Frostbelt Corporations (16.50%). However, just as we suggested, Sunbelt corporations *also* gave more of their money to the *opponents* of New Right candidates (9.12% vs. 8.04% for Frostbelt corporations). Looking at giving to one side is not enough; both sides of these races must be incorporated into the analysis.

To assess the political partisanship of corporations, a more appropriate comparison would be: What percentage of the money that corporations gave to these races went to New Right candidates and what percentage went to their opponents? That is, to what degree did each corporation show a preference for the conservative in these races? We also ran this analysis. In 1982, Sunbelt corporations are slightly more conservative, but the differences are *not* statistically significant (at the $P < .05$ level). Sunbelt corporations gave 74.95% of their money to these races to the New Right candidate (25.05% to the opponents), while Frostbelt corpora-

tions gave 72.88% (27.12% to opponents), a difference of 2.07%. After we introduce appropriate controls, Burris's principal finding is no longer statistically significant, even though we used Burris's chosen year (1982), Burris's sample (all top 1,000 corporations with PACs that contributed $5,000 or more), and his definition of the New Right.

Conclusion

In most published scholarly exchanges the person writing the comment alleges the author of the original article has made errors. For the most part that is not the case here; Burris's comment focuses primarily on differences from one election to the next, an issue that has been the focus of our own work for the past several years (Clawson and Neustadtl 1987; Su 1989; Su, Clawson, and Neustadtl 1988; Clawson and Su 1990). We encourage Burris to publish his work on other years and to provide the information needed to allow others to evaluate and replicate his analysis.

DAN CLAWSON

University of Massachusetts at Amherst

ALAN NEUSTADTL

University of Maryland

REFERENCES

Blalock, Hubert S. 1979. *Social Statistics,* 2d ed. New York. McGraw-Hill.
Burris, Val. 1987 "The Political Partisanship of American Business." *Americal Sociological Review* 52.732–44.
Clawson, Dan, Allen Kaufman, and Alan Neustadtl. 1983. "Which Class Fractions Fund the New Right?" Paper presented at the meetings of the American Sociological Association, Detroit (September).
———. 1985. "Corporate PACs for a New Pax Americana." *Insurgent Sociologist* 13, nos. 1–2 (Summer–Fall). 63–77.
Clawson, Dan, and Alan Neustadtl. 1987. "Business Elites and Political Realignment." Grant proposal to the National Science Foundation. (Available from the authors.)
Clawson, Dan, Alan Neustadtl, and James Bearden. 1986. "The Logic of Business Unity: Corporate Contributions to the 1980 Congressional Elections." *American Sociological Review* 51:797–811.
Clawson, Dan, and Tie-ting Su. 1990. "Is 1980 Special? A Comparison of 1980 and 1986 Corporate PAC Contributions." *Sociological Quarterly* 31:371–87
Su, Tie-ting.1989. "Dynamic Structure of Corporate Political Action: A Study of Corporate Political Action Committees from 1978 to 1986." Ph.D dissertation. University of Massachusetts at Amherst, Department of Sociology.
Su, Tie-ting, Dan Clawson, and Alan Neustadtl. 1988. "A Dynamic Analysis of Corporate PAC Groupings, 1978–1986 " Paper presented at meetings of the American Sociological Association, Atlanta (August).

Review Essay: The Alchemist of Contingency Theory[1]

John F. Padgett
University of Chicago

Sociological organization theory today is a mature subfield, in both the good and the bad senses of that term. There is a clear set of well-articulated contending perspectives—transaction-cost economics (Williamson 1975), resource dependency (Pfeffer and Salancik 1978), population ecology (Hannan and Freeman 1977), and neo-institutionalism (Meyer and Rowan 1977; Powell and DiMaggio 1991)—all contending over the turf of "organizations and their environments." All of these labels are battle flags coding not just particular books but entire mini-schools, each with its reasonable claim to the Lakatos nirvana of vibrant empirical research programme, rooted in well-defined assumptions and internally generated puzzles. All have a healthy dose of imperialist vision, albeit with an equally healthy low capacity for attaining it. (The situation in organizational analysis becomes even more dynamic once one considers the collateral perspectives looming just over the subfield horizon—principal-agent informational economics, structuralist network analysis, phenomenological accounting, and the "new industrial orders" of Piore and Sabel.)

Yet this apparently orgiastic panoply of theoretical choices is profoundly constraining when viewed from the perspective of the field's founders—Weber and Michels, or even Selznick and Simon. I refer of course to the disengagement of our more inwardly focused subfield from the broader study of society, politics, cognition, and control. More specifically, three institutionalized constrictions of our vision currently reign hegemonically across almost all schools of thought: (1) the concentration of "serious" organizational research in the substantive arena of economic business organizations; (2) a neo-Darwinian fixation on "fit" with "environments"; and (3) the abandonment of earlier work on microprocessual or macrohistorical mechanics, in favor of an anatomical concern with

---

[1] Address all correspondence to John F. Padgett, Department of Political Science, University of Chicago, Chicago, Illinois 60637.

structure. While radically different in significant ways, transaction-cost economics, resource dependency, and population ecology all reinforce these common premises in each other, through the medium of their mutual contention. Even the apparent exception of neo-institutionalism, I would argue, only illustrates all too clearly the hegemonist point that self-conscious struggle against dominant paradigms on their own terms only engenders (perhaps unintentionally) a securely marginalized position for oneself, while reinforcing the ongoing deep structure.

Within sociological organization theory, the point of origin for this state of affairs is contingency theory, the 1960s' progenitor of all four current contenders. Like all good ideologies, contingency theory is largely invisible as a self-conscious school today, having fractured into its more sharply focused successors. But when James D. Thompson's *Organizations in Action* (1967), Lawrence and Lorsh's *Organization and Environment* (1967), and Katz and Kahn's *The Social Psychology of Organizations* (1966) burst onto the organizational scene with their strident calls for a reorientation toward the study of "open systems" and "organizations in their environments," they single-handedly reshaped the field. Consigned to the dustheap, within sociology at least, were the earlier "unscientific" approaches of the 1950s—organizational natural histories like those of Selznick (1949) and Blau (1955); narrow-minded "closed system" ideas of Weber (1947) and the Carnegie school (March and Simon 1957); and the "overly culturalist" accounts of macrocomparativists such as Bendix (1956).

This background frames an understanding of Stinchcombe for two reasons. First, Stinchcombe's early article on the construction industry (1959) is often regarded, along with Alfred Chandler's *Strategy and Structure* (1962), as the precursor of, if not the causal agent for, the rise of contingency theory. While not a strident polemicist, Stinchcombe was very much a part of the history I have sketched. Second, however, Stinchcombe's new book, *Information and Organizations,* is an extended contemplation of and exegesis on the microfoundations of contingency theory. In this book, there is not an attempt to create a new "fifth school" of organization theory. (Not that Stinchcombe, sui generis as he is, would be capable of any such thing.) Rather, this book is a somewhat deeper effort to go back to basics and rethink the (usually implicit) microinformational mechanisms underlying the entire subfield's dominant paradigm. In terms of the hegemonic premises itemized above, Stinchcombe attempts here to explode and rework premise 3, while leaving, unexamined and intact, premises 1 and 2.

First a few words about what Stinchcombe's book is not. Architecturally, it is not really a book at all, in the narrow sense of a tightly deductive structure with clear beginning, middle, and end. The book is instead a coherent set of parallel essays thematically unified by a repeated focus on organizational structures' adapting to information needs. This architectural feature is neither to be applauded nor deplored; it simply has been and always will be Stinchcombe's style.

In addition, Stinchcombe rigorously disavows any concern with matters many would consider rather weighty—for example, power and culture. Except in a brilliant late chapter on class consciousness, Stinchcombe ascetically hews to the straight and narrow of developing a rigorously functionalist account (in the precise sense of Stinchcombe [1968]) of how organizational structures respond to uncertainty. The functional mechanism of adaptation in this book is neither self-conscious rational choice nor unintended environmental selection. Rather, for Stinchcombe, the micromechanism underlying organizational adaptation is trial-and-error learning: "The organization will continue to throw up problems to individuals until they correctly identify the source of the functional inadequacy and build a structure that remedies it" (p. 128).

This self-limiting disavowal will no doubt be a source of disappointment to some. But I regard it, along with the book's empirical attention to process and detail, as crucial to the achievement here: Stinchcombe transcends from within the contingency framework with which he begins. In this book, Art Stinchcombe is a mole burrowing his way out of the contingency theory hole he helped to create. As I shall argue below, this fact inhibits his ability to take a deep breath and survey the fresh air he breaks his way into, but Stinchcombe's achievement never would have been possible had he not been as monomaniacally committed both to consistent reasoning and to consistency with microprocessual detail.

The book itself (apart from its introduction and conclusion) contains three clusters of chapters: (1) a cluster of chapters on the microorganization of work—skill and technical information systems; (2) a cluster of chapters on governance—divisionalization and interfirm contracts; and (3) a cluster of chapters on labor selection and control—certification, class consciousness, and the special case of university personnel.

The main point of the chapters on work, derived from both Stinchcombe's earlier analysis (1959) of construction and his more recent study (with Carol Heimer) of North Sea oil (1985), is that skill is the ability to shift flexibly among an array of alternate work routines, as changing task specifications demand.[2] One (sometimes costly) way in which organizations absorb uncertainty, therefore, is to place skilled employees on the interface of those sectors of the environment where demands are most heterogeneous and shifting.

The main point of the chapters on governance is that divisionalization (or more generally "hierarchy," even when expressed in contract terms) emerges when market volatility ramifies back into manufacturing and engineering. "Markets," however, are not exogenously given and fixed by God. They are informational poolings, for both firms and consumers, of those particular bundlings of product mixes and social selling relations set up by firms' marketing departments. The chapter based on a deep rereading of Chandler's *Strategy and Structure* (1962) is one of the best

[2] Note that, for Stinchcombe, unlike Simon, skill is *not* "complexity of subroutines" (alias "chunking")—a meaning with which it is often confused.

in the book. (My only criticism is that I object when that all-pervasive word, "uncertainty," is used to cover the very different phenomena of rapid switching across different tasks or products and of time-series volatility in demand for single products. The former implies delegation of authority downward to loosely coupled artisans or professionals, while the latter evokes centralization of control within product divisions across functions.)

The main point of the chapters on labor is that segmented categorizations of labor (e.g., seniority, dual labor markets, career ladders, peer-review procedures, and even unions) are all personnel certification mechanisms that enable employers to construct, evaluate, and predict in proxy form that evanescent and often opaque property of skilled employees called "job performance." Conversely, however, such schemes also inform employees about their own interests: structural equivalence sets created by common contracts become catalyzed into class consciousness through the conjuncture of endogenous learning from incentive schemes and exogenous cultural "injustice frames." Stinchcombe sees two second-order mechanisms as available to employers for demobilizing such budding class warfare: (1) constitutionalism, wherein employers disappear behind a naturalized veil of "objective" rules, and (2) individuation, wherein employees' attention is redirected toward distinctions among themselves, either directly via competitive piece-rate schemes (Burawoy 1979) or indirectly via external peer review.

Underneath all these analyses is a vision of information as "news." That is, according to Stinchcombe, organizations do not gather information primarily in order to anticipate future states of the world rationally, through some sort of Bayesian probabilistic projection. (Once collected, however, information might indeed be used for that purpose.) Rather, organizational structures expand toward the social sources of quick-breaking *particular* outcomes in order to respond rapidly to (relative) certainties. A major complication of which Stinchcombe is well aware is the issue of buffering or pooling—that is, the level of aggregation at which information, even as news, is defined. But whatever the category schemas underlying the social construction of information, Stinchcombe's functional axiom remains.

How, by these arguments, does Stinchcombe transcend his own contingency theory framework? Where exactly in the analysis does Stinchcombe's burrowing hit fresh air? And what is the character of that fresh air that Stinchcombe himself only dimly perceives? After all, when it is stated at the level of aggregate, organizational propositions, much in this book is really old stuff. Propositions like "heterogeneous environmental demands lead to craft-style flexible specialization" and "volatile, technologically separable markets induce multidivisional forms" are not exactly news to the subfield right now.

Stinchcombe's contribution, therefore, is not primarily at this level of aggregate predictions; it is, instead, at the level of regrounding such propositions in locally lived experience, through the orchestrated elabora-

tion, across a variety of settings, of micromechanisms like trial-and-error learning and information as news. The payoff of this move is the dereification, indeed the virtual dissolution, of the concepts of "organization" and "environment." Organizational structure is still the concern, first and foremost, but instead of a botanical emphasis on the comparative anatomy of forms, we are returned to an almost old-fashioned microfocus on process and mechanism (albeit informational, rather than political, in character).

I was particularly impressed by Stinchcombe's almost unintentional deconstruction of the concept of "market," both on the side of sales and (especially) on the side of labor. On the sales side, Stinchcombe's sensitive discussion of Chandler's case study of Sears is exemplary. Whether small tractors are packaged with lawn fertilizers to make retail "garden stores," or whether all tractors are sold together in John Deere outlets makes all the difference to Stinchcombe in how markets are assembled by firms, in how choice, information, and uncertainty become represented, and in how social relations develop between customer and firm (see Douglas [1986] for an analogous example of the institutionally grounded difference between American and French classifications of wine).

On the labor side, Stinchcombe is, if anything, even more sensitive in his inductive portrait of how firms actively create labor categories and reference sets for control, even though they thereafter become caught up in the contingency theory adaptational dynamics of their own creation. I say "even more sensitive" because of the second-order mechanisms referred to above. Firms not only construct the comparison set categories of their customer and labor alters; through constitutionalism and individuation, they also manage the relations of those alters to the (now largely subliminal) categories. This style of argument, inductively driven by microdata on process, is what I mean by fresh air.

Having mostly complimented Stinchcombe so far, let me now point out that I do not believe that Stinchcombe fully appreciates the implications of what he has wrought.

Throughout most of the book, "information" has a curiously flat, inert character—nothing more than an objective fact resource to be processed, like rubber or steel. The exception is in Stinchcombe's brilliant treatment of class consciousness, in which information comes alive through workers' reflexive reshapings of their own identities. But Stinchcombe never does the same thing for managers, even though March and Simon (1957) long ago pounded on the point that organizational information structures shape the "premises of decision" of everyone in organizations. The point could be extended to consumers as well. What is missing, in other words, is some sense of the agency of information—how information flow redounds on *both* sides of the transaction to reshape actors' inferences about the conceptual constituents of their own worlds and about their own interests, in situ. Stinchcombe should, and can, take more seriously the fact that information does not just passively measure the movements

of objects in the world; it also induces users to adopt latent definitions of units and actors, through arranging sets of "comparables" within which objects are perceived.

To show how such an extension would indeed be consistent with his overall project, let me close with some illustrative reworkings of Stinchcombe's material. In particular, I shall concentrate here on three archetypal modes of classificatory representation, which I (and probably Stinchcombe) believe underlie firms' informational constructions of both labor and consumer markets: (a) broad homogeneous products/jobs, within which is induced competition over quantity; (b) stratified arrays of differentiated product lines/careers, within which is induced contested mobility within channels; and (c) bounded demarcations of customized services, within which reputation, honor, and an individuating mystique of quality and autonomy are used for control.

Competitive piece rates (Burawoy 1979) and Fordist mass production (Sabel 1982) are examples of the first mode of firms' classificatory constructions of workers and consumers, respectively. Informational aggregation—that is, grouping heterogeneous others into broad, standardized categories—permits managers to remove themselves from the fray by transforming potentially hostile others into "law of large numbers" natural laws. Indeed, rational planning and forecasting by managers are triggered by this. Organizational adaptation in this mode becomes largely a matter of time-series volatility: vertical integration or "tight-coupling" across functions is inversely related to how well pooling (of orders) and buffering (of inventories) arrangements can be stabilized.

Whether or not any mode of collective representation locks in for profits and for "naturalized control," however, depends on whether or not the workers or consumers so classified can be induced into competitive "deep play" with each other. In this first mode, deep play takes the form of intense status competition among directly interacting others, sociologically similar to each other. Burawoy's (1979) example is the machine-tool workers' "making out" shop-floor culture, in which sharp but potentially fluid pecking orders generated by workers discipline and mobilize workers' identities for the firm. Neighbors locked into suburban tract housing come to mind as a consumer-side analogue (as do the comparisons my colleagues and I make over lunch about our latest PCs). In all these examples, quantitative add-ons onto a standardized (but hopefully, from the firms' point of view, changing over time) base is the key to successfully inducing status anxiety within role sets.

But standardized markets constrain firms as well as workers and consumers. For firms, a common route out of such markets is product differentiation—the joint slicing up of broad categories into more specialized niches within which local monopolies obtain, or, even better (since specialization is risky in the face of volatility), the construction of whole arrays of specialized products (i.e., "lines") through which consumers are induced to step. The 1920s victory of General Motors over Ford is the classic example. But the packaging of fashions into Perry Ellis and

Calvin Klein brand name "looks" (further assembled into Nordstrom arrays) is the same thing.

The basic ideas in this second mode of consumer/worker categorization are hierarchy and turnover. Lines are anchored by the highest status products, but consumers are given narrowly tracked, step-by-step ladders of related products through which to scale the heights. Deep-play identity here is the aspired, not the achieved, whole line; in this way, consumers' status anxiety gazes are refocused from peers onto superiors. Turnover is central, since for firms profits come from consumers climbing the whole ladder, not from their jumping immediately to the top.

The obvious analogue on the labor side is the organizational career— the stepwise arrangement of differentiated jobs into clearly verticalized arrays. Difficult but eventual mobility is the key for maintaining deep engagement by workers; the idealized arrangement of mobility into clear tracks, with visible yet distant terminal nodes, is the key to keeping such engagement focused on the firm. Finely grading the ladders is one mechanism for keeping mobility rapid yet rates of progress low.

In both the labor and the consumer versions of this control mechanism, the task of the manager is not the rational projection of cost and demand; it is the arrangement of jobs and products into coherent and self-sustaining differentiated arrays. (At the highest conglomerate level, indeed, "strategic planning" means arranging portfolios of markets within firms; see Eccles and White 1986.) The goal in these efforts is not the maximization of transactional profits: market rigors are transcended through long-run identification of workers and consumers with firms. The irony, of course, is that potent firm control (via the mechanisms of identification and of perceived interdependencies and sunk costs) is achieved through the presentation of an apparent orgy of choice: tightly orchestrated arrays, after all, contain many steps.

The final mode of consumer/worker representation I draw from Stinchcombe is individuation. This mode has long engaged Stinchcombe's attention, of course, under the label of "customized craft production." Stinchcombe's extrapolation of this mode to the realm of university professors in this book, however, has the effect of raising Stinchcombe's (and our) vision from the level of the microorganization of work to the level of how personnel are evaluated and jobs distributed within an apparently individuated world of idiosyncratic products, flexible specialization, and "autonomous" artisans/professionals. Stinchcombe's key point (familiar to anyone who has gone through tenure review) is that extrafirm reputational networks of invisible college control underpin apparent local freedom and craft autonomy. "Do your own thing" and "Be flexibly specialized" says the tantalizing official rhetoric; but wise craftsmen never forget that hidden elite brokers code their "free" actions into reputation, and control thereby all future jobs and rewards that flow their way. (Stinchcombe never brings this point back to his original world of housing construction, but clearly it is as true there as it is in production realms as disparate as doctor referrals and Third Italy textiles.)

The consumer version of this individuation mode of informational control should be familiar to anyone who has ever been in a hospital or had his or her car repaired. "Relational contracting" is the face-saving art of negotiating the terms upon which you abjectly surrender, all in the names of personalized service and of a paternalistic concern for your welfare. (More "serious" industrial examples include U.S. government purchases of defense armaments and Japanese suppliers of parts for Toyota.) Instead of aggregating customers through statistics, customers and/ or tasks are fragmented into "incommensurate" cases. Subtler modes of indirect control through the redirection of customers' gazes are eschewed for direct "bear hug" domination based on extreme asymmetry in information flow.

Labor and customer forms of individuated control, despite appearances, are actually the same because behind particularistic processing of jobs and people lies a private coding of whole persons' relative honor or worth. Precisely because codings emerge through informal networks of conversations among specialists, the logic of the system appears (perhaps blithely) opaque in the eyes of workers or consumers. Needless to say, intense degrees of structural concentration and inbreeding can develop within such reputation-coding informal networks (Faulkner 1983).

Extensions aside, Stinchcombe's *Information and Organizations* will be a core book for organization theorists interested in the foundations of their subfield. The book could easily be misread as just another permutation of contingency theory. However, implanted in Stinchcombe's inductive style is a revolutionary transformation that deserves wide thought.

REFERENCES

Bendix, Reinhard. 1956. *Work and Authority in Industry*. New York: Wiley.
Blau, Peter M 1955. *The Dynamics of Bureaucracy*. Chicago: University of Chicago Press.
Burawoy, Michael. 1979 *Manufacturing Consent*. Chicago: University of Chicago Press.
Chandler, Alfred D. 1962. *Strategy and Structure*. Cambridge, Mass · M.I.T. Press.
Douglas, Mary. 1986. *How Institutions Think*. Syracuse, N.Y.· Syracuse University.
Eccles, Robert G., and Harrison C. White 1986 "Firm and Market Interfaces of Profit Center Control." Pp. 203–20 in *Approaches to Social Theory*, edited by Siegwart Lindenberg, James S. Coleman, and Stefan Nowak. New York· Russell Sage Foundation.
Faulkner, Robert R 1983 *Music on Demand: Composers and Careers in the Hollywood Film Industry* New Brunswick, N.J : Transaction Books.
Hannan, Michael T., and John Freeman. 1977. "The Population Ecology of Organizations." *American Journal of Sociology* 82·929–64
Katz, Daniel, and Robert Kahn. 1966. *The Social Psychology of Organizations*. New York: Wiley.
Lawrence, Paul R., and Jay W. Lorsch. 1967. *Organization and Environment*. Cambridge, Mass.. Harvard University Press
March, James G., and Herbert A Simon 1958 *Organizations*. New York: Wiley.

Meyer, John W , and Brian Rowan 1977. "Institutionalized Organizations· Formal Structure as Myth and Ceremony " *American Journal of Sociology* 83.340–63

Pfeffer, Jeffrey, and Gerald R. Salancik 1978. *The External Control of Organizations* New York. Harper & Row.

Powell, Walter W., and Paul J. DiMaggio. 1991 *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.

Sabel, Charles F. 1982. *Work and Politics: The Division of Labor in Industry*. Cambridge: Cambridge University Press.

Selznick, Philip. 1949. *TVA and the Grass Roots*. Berkeley, Calif. University of California Press.

Stinchcombe, Arthur L. 1959 "Bureaucratic and Craft Administration of Production." *Administrative Science Quarterly* 4.168–87.

———— 1968. *Constructing Social Theories* New York. Harcourt, Brace & World

Stinchcombe, Arthur L., and Carol A Heimer. 1985. *Organization Theory and Project Management: Administering Uncertainty in Norwegian Offshore Oil*. Oslo Norwegian University Press.

Thompson, James D 1967. *Organizations in Action*. New York: McGraw-Hill.

Weber, Max. 1947. *The Theory of Social and Economic Organization*, edited by Talcott Parsons. New York: Free Press

Williamson, Oliver E. 1975. *Markets and Hierarchies* New York: Free Press.

Book Reviews

Freedom in the Making of Western Culture. By Orlando Patterson. New York: Basic Books, 1991. Pp. xviii + 487. $29.95.

Martin Bernal
Cornell University

Freedom in the Making of Western Culture is an investigation of the well-known hypothesis that there is a correlation between the rise of the concept of freedom and the existence of slavery. Orlando Patterson approached this topic as a distinguished student of slavery in both the ancient and modern worlds. This volume, however, is not concerned with the modern world but with what the author sees as the development of the concept of freedom in ancient Greece and Rome and in early Christianity with a relatively short discussion of the Middle Ages. He justifies this focus on the reasonable grounds that the concept of "freedom" was fully worked out before the rise of capitalism. The author's scholarship and coverage of the literature of modern classics are outstanding and these, combined with his vantage point as an outsider, have allowed him to create a massive synthetic work on a scale that would have been impossible for someone enmeshed in the relevant fields. He has also used his sensitivity to the pains of slavery and the situation of women to write some impressive chapters that go beyond conventional scholarship concerning the centrality of women's longing for freedom to Greek tragedy.

Nevertheless, I believe that the work is fatally flawed. This is because the author has failed to transcend the secondary literature on which it is based. Patterson has accepted the "Western" monopoly of freedom and of the "Greek miracle" in both the spirit and the letter. The adjective "Western" is used in tones of awe and everything non-Western is treated with hostility or condescension. The first section is entitled "The Stillbirth of Freedom in the Non-Western World." He begins the subsection "Slavery and Conceptions of Freedom in Preliterate Cultures" with references to ancient China and Carthage, societies that were, in fact, highly literate. As it so happens, Carthage was also a "slave society" with a considerable concern for citizens' freedom. Incidentally, the conjunction of the two shows Patterson's Eurocentric inclination to lump extraordinarily different societies under the single heading of "Non-Western World."

My objections here are not merely based on a vague sense of "political

correctness." I believe that wearing these Eurocentric blinkers makes it impossible for Patterson and the scholars upon whom his work is based to explain or even to describe the origins of chattel slavery and freedom. One cannot simply look at Greece in isolation; it is necessary to see it in its geographical and historical context. According to Patterson, slave society (one in which the proportion of slaves in the population was large enough to make the demarcation between slave and free central to the consciousness of society as a whole) emerged, without outside stimulus, in Greece in the 7th century B.C.

Patterson claims that the "primitive democracy" found in ancient Mesopotamia, whose existence has been demonstrated by Thorkild Jacobsen, could not be sustained in larger and more elaborate societies and that the Greeks were the first to combine mass participation with large-scale government. This is nonsense. While it is true that, as in ancient Greece and modern Europe, there were strong autocratic tendencies in ancient Southwest Asia, there were also continuing constitutional and participatory traditions. Just to take two examples, let us look at the Syrian cities of Ebla in the third millennium B.C. and Ugarit in the second millennium with populations of tens if not hundreds of thousands and complicated trading and international relations. Both cities had intricate constitutions of a generally "Venetian" type in which the "king" was in many ways similar to the doge and dependent on other mercantile families. In them, as in most Near Eastern cities, there appear to have been regular meetings of elders and wider assemblies of the male citizens. At times in certain cities, these assemblies appear to have been the supreme power. In the 14th century B.C., the Egyptian pharaoh addressed "the People of Arwad" as the "sovereign" of that important and independent coastal city. There was also considerable servitude in Mesopotamia and Syria.

Patterson rightly points out that it is not useful to call either Mesopotamia or Egypt slave societies. However, this is a good description of Phoenicia from the 11th century B.C. In this century, the Phoenician ports developed into a new kind of city-state, with complex, varied, and shifting constitutions, some involving the assembly of citizens, drawn from their own and Mesopotamian participatory civic traditions. However, while the economy of the latter was founded on agriculture those of the Phoenician cities were based on trade and manufacture, much of which was carried out by chattel slaves. All the biblical and early Greek references to Phoenicians mentioned slaves. The cultural centrality of the master-slave relationship in the culture of Phoenicia and its neighbors is shown by its use in Israel—where there were relatively few chattel slaves—as a metaphor for God as a slavemaster who sells and redeems his people. In becoming a "slave society" Phoenicia appears to have made liberty a central value. There is no doubt that a passion for freedom runs through the Old Testament, and the Hebrew and presumably the Phoenician word for "freedom," $d^e r \hat o r$, has the poetic images of free-flowing rivers and swallows that never touch the ground. Hence, it is

hard to see on what grounds the earlier West Semitic concept should be placed on a lower level than the Greek *eleutheria* or the Latin *libertas*.

There is considerable documentary and archaeological evidence indicating contact between Phoenicia and the Aegean between 900 B.C. and 700 B.C. This makes it difficult to maintain that the emergence of similar city-states, a slave society and its concomitant concept of freedom, in Greece in the 8th and 7th centuries B.C. were independent developments. Patterson maintains that Greece created slave society and invented "freedom," ex nihilo. Such a position was always implausible. Given the advances of Near Eastern scholarship, it is now untenable.

The question of whether or not there is a continuous Western tradition of freedom from Athens to America is too complex to go into in this review. However, the case for it is not made persuasively in this book. Thus, I am not convinced either that the classical Greeks invented freedom or that there is one Western thread leading from ancient Greece to the modern world. There are many more discontinuities and reinventions and the game had many more players than he supposes. I am sad to say this about such a distinguished sociologist and authority on modern slavery, but, in this work, I see him as a good man fallen among the narrow scholars he thanks in his preface.

Minds Stayed on Freedom: The Civil Rights Struggle in the Rural South—an Oral History. By the Youth of the Rural Organizing and Cultural Center. Boulder, Colo.: Westview Press, 1991. Pp. vii + 189. $37.50 (cloth); $12.95 (paper).

Aldon Morris
Northwestern University

Minds Stayed on Freedom is a unique oral history of how the civil rights movement developed and was sustained in a rural county in Mississippi. The experiences of the poor and disenfranchised masses are almost exclusively interpreted and written about by people who are not members of those populations. Indeed, writing and publishing are dominated by privileged members of a society who are usually college educated and well-connected to the organs of the cultural aristocracy. This practice is so routine that it has become a taken-for-granted reality. In America, it is commonplace for elite whites to virtually control the written word when it comes to interpreting the experiences of African-Americans. One only has to peruse the *New York Times Book Review* to see that the overwhelming majority of books on the African-American experience are reviewed by whites.

Minds Stayed on Freedom suggests that poor and culturally disenfranchised people are fully capable of interpreting their own experiences from the bottom up. This oral history of the civil rights movement in Holmes County, Mississippi, was researched and written by 18 black, eighth and

ninth graders associated with the Rural Organizing and Cultural Center located in the same county. These youths conducted interviews with some of the key persons who founded and sustained the Holmes County civil rights movement. They also transcribed and edited these interviews and wrote useful vignettes about each of the interviewees.

What emerges from this work are powerful accounts of how this significant movement unfolded in a rural southern community characterized by vicious white oppression and the absence of national media attention and famous charismatic leaders. Poor black farmers took the lead in developing this movement. Indeed, from these accounts, it is clear that black land ownership was a critical factor in determining who would develop and lead the movement because such ownership provided these individuals with the independence needed to challenge white domination. Members of the black middle class—teachers and preachers—usually joined the movement after much of the dangerous work had been done by poor farmers, given that whites controlled the entire school system and targeted potential activist churches for bombing raids. This finding flies in the face of those studies that claim the civil rights movement was a black middle-class movement unconcerned with the interests of the black poor and working class. We also learn that black women in the rural South were central to both the leadership and grassroots organizing of this movement. It becomes clear also that the participants' identification with the cause of black liberation was so strong that it generated a commitment that could not be silenced by the fear of crosses burned by hooded Klansmen, the prospect that one's home or church would be blown asunder by dynamite, or the view that one's own life was more important than liberation. Social-movement scholars have come nowhere close to understanding this phenomenon and certainly rational choice models fail to explain why some people would rather sacrifice their lives than tolerate oppression for future generations. Finally, this work also makes it clear that many civil rights participants in the rural South employed nonviolence as a tactic while they guarded their homes, churches, and loved ones with loaded shotguns and Saturday night specials.

Thus, the young authors are to be commended for the fine interviews they conducted with their elders on this important movement. They have done a valuable service for movement scholars and for those who would struggle in the trenches to foster social change. In addition, Jay MacLeod has written a useful introduction to the volume. He provides a history of black/white relations in Holmes County, Mississippi, and he situates this rural social movement within the national civil rights movement. In haste to differentiate this rural movement from the larger movement, he overstates the differences. A few examples will illustrate the point. Much is made of how this movement relied on armed self-defense rather than Martin Luther King's nonviolent strategy. Yet, the data from the interviews stress the importance of nonviolence as a strategy. Moreover, self-defense of the sort that was utilized in Holmes County was widespread throughout the larger movement. The centrality of the black church is

somewhat underemphasized. Many of the citizenship classes where consciousness raising and strategizing occurred were centered in the church. The symbolic inner life of the people who staged the confrontations were anchored in black religion and the church. Thus one of the interviewees explained that she was able to stare danger in the face "because I know I had somebody there who was on my side. And that was Jesus; he was able to take care of me. That who I depend on and put my trust in" (p. 138). The importance of the black church in the origins and development of the civil rights movement was both organizational and deeply mental. Nevertheless, exaggerations not withstanding, MacLeod's introduction successfully situates the Holmes County movement within its larger historical and social contexts.

In short, this book will be useful for both movement scholars and students interested in social movements and social change. It is also proof that people on the bottom can think, write, and act.

In Their Own Interests: Race, Class, and Power in 20th-Century Norfolk, Virginia. By Earl Lewis. Berkeley and Los Angeles: University of California Press, 1991. Pp. xv + 270. $34.95.

Rutledge M. Dennis
George Mason University

In Their Own Interests is a major contribution toward understanding the political economy and culture of an urban center in the upper South, namely, Norfolk, Virginia. More specifically, it is a closely woven account of the making of black Norfolk, of the factors that precipitated the emergence of an urban industrial class and a politically active middle class, and of the changing role of race and class in Norfolk. Focusing on the years 1862–1945, Lewis narrates a lively and vivid history of how African-Americans in Norfolk forged links between the workplace and the home sphere. These two concepts constitute one of the major themes of the book. Lewis defines the workplace as the larger world beyond the black community. According to Lewis, this world was larger than the workplace because, in addition to including the fight for employment, equal wages, and union membership, it also included the fight for a "greater voice in social and political affairs" (p. 163); the home sphere includes an area larger than the home, for it entails "improvements within the community, neighborhood, and church and the shared culture and expectations" (p. 206) forged among friends and relatives that help to create feelings of collective unity. Lewis cogently analyzes the workplace and the home sphere as dichotomous and dialectical, explaining how the exigencies of race and class forced blacks to move consciously between the two poles, now shifting toward the workplace, later moving toward the home sphere, at other times moving simultaneously toward both poles.

Another major theme permeating the book is the dynamic intersection of class and race relations in Norfolk. Citing previous class/race analysis as rather rigid, inadequate, and simplistic, Lewis explains that the race/class issue is best understood along the dialectical lines he uses in assessing the workplace/home sphere dichotomy. According to Lewis, race and class consciousness among a well-informed and growing middle class and a rather activist labor class in black Norfolk forged a unique understanding of the role of race and class in the racial politics of the city. This awareness made it possible for blacks to construct a variety of tactics and strategies and to shift between the two poles whenever their interests dictated the need for such shifts. For example, Lewis notes the importance of class when he presents the history of attempts by Socialists and Communists to establish the idea of workers' solidarity among blacks and whites during the late 1920s and throughout the 1930s. The race theme is highlighted in the author's discussion of the emergence of the Garvey Movement. It is important to note that the Garveyites did not initiate the idea of self-help or institution building in the community. Rather, they accelerated already extant movements in both. It is on this issue that Lewis has provided a valuable service. He dug into the local history and has now alerted us to the fact that the Garvey Movement may have had a greater impact on southern urban history than is generally perceived. In any case, the call for class solidarity by the Socialists and Communists and the urgency of the need for racial solidarity as seen by the Garveyites and the legion of Garvey supporters and sympathizers determined how blacks responded to such issues as joining all-white organizations or creating all-black ones and engaging in or opposing limited alliances with whites on a variety of issues. But, as Lewis assures us, working the dialectics was not always easy and problem free; there were inherent contradictions and ambivalences in the very process itself. In a sense, blacks were entangled in a "double paradox," that of dealing with the dialectics of class and race while simultaneously juggling the dialectics of the workplace and the home sphere. Lewis states that the key to the success of blacks was framing both dialectics around the issue of power and power inequality between blacks and whites, not only in Norfolk but throughout the country. And because the issue was power, blacks had to wage a battle on several fronts: to fight for "interracial solidarity in the workplace and for ethnic and racial solidarity in the quest for community empowerment" (p. 188).

Confronted with the problem of having to live, as DuBois called it, "behind the veil," Lewis documents how blacks shaped a life of their own behind their own veil, a life removed from whites that demonstrated that distance and separation imposed by whites would guarantee blacks some degree of individual and community empowerment. On the question of empowerment Lewis cites the important role of indigenous black organizations and institutions such as the church, social and civic groups, sororities and fraternities, and the local all-black unions in creating and sustaining the social and cultural values that promoted and legitimized

the concept of empowerment. In analyzing the relationships between indigenous institutions, social values, and black culture, Lewis, like Ralph Ellison, acknowledges that social and cultural values in the black community did not develop in a vacuum, but that many of the components of black culture may be understood within the contexts of the social and institutional networks that blacks themselves have woven. But having made a case for viewing community cultural and social values semiautonomously, Lewis hastens to add that it is an error to view the black community as "monolithic." Rather, an accurate assessment of the community would entail acknowledging much diversity and how such diversity plays itself out in the degree of tension and conflict between groups. He presents examples of conflicts between conservatives and radicals and between males and females. Lewis characterized the conservatives as continuing to believe in the traditional value of racial paternalism and racial "go-slowism." The radicals consisted of those, generally younger, who believed in a rapid restructuring of the system in order to increase black participation in all spheres. But the irony of success for the radicals meant, paradoxically, "that greater political participation meant the death of the race exclusive organizations which had benefitted blacks well over the years" (p. 197).

Lewis's probe into male-female tensions and conflicts is not extensive yet does provide us with some of the causes and consequences of black male-female tensions and conflicts. He notes that some of the reasons for the conflicts centered around different perceptions about work, decision making at home, and general power relations between the sexes.

Throughout the book Lewis documents how blacks were able to fight and challenge but also how they constructed a way of life based on their perceptions of their past and what they hoped would be their future. The key to all of this, according to Lewis, was a sustaining collective ability to shift their positions and to move into new and different gears as political and economic situations dictated. As Lewis states, "throughout the American Experience, the feature which enabled blacks to survive is the creativity of Afro-Americans during times of crisis, the restructuring of social relations, and a rethinking of how best to act in their own interests" (p. 126).

This is an excellent account of the social history of a community and the political economy and culture that shaped the contours of black and white Norfolk. It is well written and very insightful. It is also a pleasure to read the social history of a community by a historian who is also sociologically informed. Lewis's detailed analysis of family and individual life histories and the manner in which he juxtaposed these histories against the larger community issues and concerns made his characters more lively and livable. In that sense, he was an adept urban anthropologist weaving family, individual, and community into a social web or network. The book is largely free of academic jargon, is very readable, and should be of interest to both social scientists and the lay public.

Indian Country, L.A.: Maintaining Ethnic Community in Complex Society. By Joan Weibel-Orlando. Urbana: University of Illinois Press, 1991. Pp. xii + 354. $34.95.

C. Matthew Snipp
University of Wisconsin—Madison

In the 1950s and 1960s, American Indians moved *en masse* to urban areas, partly because federal authorities encouraged them to do so. The urbanization of American Indians captured the attention of social scientists who produced numerous studies on this subject in the 1960s and 1970s. Since then, little has been written about urban Indian social life.

This paucity of research makes Joan Weibel-Orlando's *Indian Country, L.A.* an especially timely contribution. It provides a remarkably in-depth examination of the Los Angeles Indian community. The choice of Los Angeles is particularly auspicious because this city has the largest, most diverse urban Indian population in the nation.

Indian Country, L.A. is a large book: it contains 16 chapters organized in six sections. However, it actually consists of three parts. The first part is background material, the second part discusses roles and institutions, and the third part is a case study of a community crisis.

The background material presented in the first part of the book covers a wide range of theoretical and methodological issues. As an anthropologist, Weibel-Orlando seems obliged to begin with a discussion of historic Indian settlement patterns in the Los Angeles basin, along with contemporary demographic information. Methodologically, her work is based on participant observation and in-depth interviews with key informants. Theoretically, she takes up issues such as ethnic identity and the meaning of community. While all of this material is relevant to her study, I found this part of the book to be the least engaging. The exception is a chapter on American ethnic identity that is extremely insightful.

Weibel-Orlando's analysis of roles and institutions is truly an excellent contribution to writing about American Indians. To capture the diversity of the Los Angeles Indian community, she focuses on three arenas of social interaction: pow-wows, gospel singings, and community service organizations. Many social scientists are familiar with pow-wows and community service organizations. Pow-wows are community celebrations with native music and dance, while the organizations typically provide assistance to needy members of the community.

However, Weibel-Orlando deserves special credit for writing knowledgably about gospel sings and the role of the church within the Los Angeles Indian community. Virtually no one has written about this subject and certainly not in as much detail. As it happens, like many African-American communities, the church and gospel singing have a prominent role in the lives of many American Indians, especially those from eastern Oklahoma. Weibel-Orlando explains how this tradition was

transplanted to Los Angeles from rural Oklahoma by immigrant Choctaws.

To illustrate roles within these community institutions, she focuses on exemplary individuals within each institution she discusses. Her actors include a church elder, an active pow-wow participant, and the administrator of a community service organization.

The last part of *Indian Country, L.A.* deals with the community response to the closing of the Los Angeles Indian Center. The Los Angeles Indian Center was once a large multimillion-dollar social service organization that served as a hub for the L.A. Indian community. Fiscal mismanagement of several large grants caused its closing. Weibel-Orlando discusses the impact of its closing and how other community groups intervened to fill the void that resulted. The purpose of this analysis is to demonstrate the cohesiveness of the Los Angeles Indian community in a time of crisis.

The only serious flaw in this book is that it sometimes strains for scientific credibility by substituting intellectual jargon for readable prose. One example. "The frontiers at which two culture-specific administrative styles converge can be the locus for positive creative innovation and synergistic growth and for negative exploitation of ambiguity" (p. 282). There is enough of this kind of writing to be annoying but fortunately not enough to be a serious problem. This is a praiseworthy book that I would recommend to a wide variety of readers: students and colleagues with interests in community, race and ethnicity, and of course, American Indians.

The Feminization of Poverty: Only in America? Edited by Gertrude Schaffner Goldberg and Eleanor Kremen. New York: Praeger, 1990. Pp. xii + 231. $45.00 (cloth); $15.95 (paper).

Diana M. Pearce
Stanford University

Is the feminization of poverty a phenomenon limited to the United States? That question is both the title and the starting point of this book, but the authors quickly broaden their inquiry to women's inequality, which is primarily economic, but is also social and political.

The authors compare the United States to six other countries that are also industrialized countries, but with different histories, racial/ethnic makeups, and political systems. Canada, Sweden, France, Poland, European Russia, and Japan. The discussion in each country chapter is loosely organized around four "factors": labor-force status of women (participation rates, part-time work, occupational segregation, etc.), policies to promote women's labor-market equality, social-welfare benefits, and de-

mographic factors. In addition, there is (usually) a discussion of the dual role of women, the situation of single mothers, and women's political power.

In a field—women's poverty—in which very little is written, even about the United States, this book is a gold mine of information. Particularly for researchers and advocates who have pushed for particular benefits or laws, it is heartening to discover, for example, that Canada has had paid maternity leave for 20 years (under its unemployment insurance system), and that 97% of families in Sweden live in fully adequate housing *and* pay an average of only 20% of their income for housing. Indeed, the pages of this book contain a rich agenda of potential policies to combat women's poverty and inequality, such as the Swedish provision for parents to take child-care leave or Japan's menstrual leave. At the same time, the loose organization and inconsistent presentation requires real digging, as in a mine, to find specific information on a topic. Even the basic question posed by the title—whether the feminization of poverty exists outside America—is answered for each country differently, using different criteria and facts. Within chapters, moreover, there is an annoying tendency to return to a topic several times, rather than to cover it at one sitting (see, e.g., divorce in Japan).

Each chapter is written by a different author or authors, with the editors contributing several chapters. The individual authors clearly have access to primary language sources, and/or documents or sources that would be difficult or impossible to access, resulting in a rich and informative perspective on women's status in each of these countries. The various authors also provide a holistic view of each society without an artificial forcing into boxes; thus the pronatalist policies that are anomalies (higher benefits for third and fourth children), and would be missed in a rigid comparative scheme, are put into context here. At the same time, the varied perspectives are frustrating. Unlike the Kamerman and Kahn studies, this study does not attempt to provide comparable figures *across* countries, even on divorce rates, birth rates, or proportion of women working part-time. Even when they are available, and at least roughly comparable, such figures are not given, one longs for a few comparative charts in the conclusions.

The dominant impression left the reader is one of having encountered a mass of "factoids"—discrete bits of information, as in the Harper's index. As at a smorgasbord, one comes away stuffed, but without a clear sense of what one has ingested. There is far too little analysis of the findings, either at the country level or comparatively. Yet provocative comparisons and implicit generalizations abound: Canada's teen pregnancy rate is similar to that in the United States, but without the racial aspects—challenging America-centered theories. One author's observation that social-welfare expenditures were higher in nations with more centralized governments raises the question whether women, particularly single mothers, will be poorer in a decentralized (or broken up) Soviet

Union, or indeed, a post-Reagan United States? Is part-time work inevitably converted in postindustrial economies into a ghetto of low-wage, temporary jobs held disproportionately by women?

Probably the weakest aspect of the book is its coverage of racial/ethnic issues. In spite of making an excellent case for doing so, the best country discussion—the United States—did not mention Hispanics, Asians, or American Indians, while other countries got even shorter shrift (two paragraphs on Sweden, where one-eighth of the population is immigrant/minority, and none on Japan's minorities).

In spite of its shortcomings, this book provides much that is missing in the literature on women or poverty. It gives a broad overview of demographic trends and women's economic status—especially employment—set in the context of each country's unique political and cultural history. Though poverty is not central to the discussion—reflecting severe data limitations as much as authors' choices—the approach is directly relevant to understanding poverty. As with trips abroad, this survey returns the reader to his or her own country with a fresh perspective on the ongoing tragedy of the feminization of poverty.

Shifts in the Welfare Mix: Their Impact on Work, Social Services and Welfare Policies. Edited by A. Evers and H. Wintersberger. Boulder, Colo.: Westview Press, 1990. Pp. 412. $38.00.

Jerald Hage
University of Maryland at College Park

The mounting crises in the welfare state in Europe are leading to a series of conferences about these problems. *Shifts in the Welfare Mix* is one example and it suffers from many of the usual problems of conference books. Before providing a general critique, I want to review some of the themes in what is, in many ways, a diverse set of papers. Given the limited amount of space, I can only mention a few of the chapters by name. I concentrate on those that offer the most insights.

The book does not start well because, while the introduction, "Shifts in the Welfare Mix," by A. Evers, attempts to place the variety of experiments or ways in which welfare services can be provided in a general context by suggesting a welfare triangle of the market economy, the state, and households, it does not locate this in the general literature on governance. Nor is the analytical problem of why there is crisis posed.

The best single country studies are those on Sweden, France, Germany, and England. Each of them attempts to provide some analysis. The best of these is the one on Sweden. Since it was written before the recent elections in that country, the article is timely because it suggests

the growing issue of legitimacy. This chapter is also good because it provides several interesting and provocative case studies.

The other chapters on Western Europe either attempt to provide too much information in too short a space or else do not place the societal experiments that are occurring in some context, either historical or sociological. None of the articles provide the reader with any sense of how typical these experiments are or whether they are the wave of the future. And this is the major limitation of the book: its lack of analysis of why particular shifts occur in specific historical contexts.

There is also a section of four chapters on Eastern Europe, more specifically Hungary, Poland, and Yugoslavia. E. Sik and I. Svetlik attempt to provide an introduction entitled "Similarities and Differences" but miss the crucial fact that in the area of welfare, East and West are much more alike than either is like the United States. The authors fail to understand the role of planning in France or the corporatist arrangements in most of Western Europe, except for Britain, which is again a similarity. They are quite right in underlining the special role of the business firm in providing welfare services, however, in the eastern part of Europe.

The two chapters on Hungary and Poland reflect the same problems described above. However, the chapter on Yugoslavia places the experiments in the larger context as attempts to market some services, rationalize others, and externalize still others. Because this provides a more general theoretical perspective, it is much more interesting.

The last section contains two articles. The most exciting is by S. Miller on the "Evolving Welfare State Mixes." It places a number of problems of the welfare state in a comparative context and observes that there are four different models of how the welfare state can change. The book concludes with a general statement by A. Evers and H. Wintersberger.

The book contains many easy-to-present omissions that could have provided a way to integrate the disparate country studies. First, we are not told how much of a crisis the welfare state is in. Presumedly, and indeed, this is the case, the extent of the crisis varies from one society to the next. Second, there should be some agreement about the scope of welfare. Some articles include education, an arena concerned with upward social mobility rather than income maintenance, and the connection between education and unemployment is never discussed. Third, welfare systems are not placed in societal contexts that would help the reader appreciate the similarities and differences between them. For example, we should be given some idea of the relative importance of the voluntary sector or the third sector, which can be quite substantial, even in societies with strong central state control.

Surprisingly most of the chapters focus on unemployment as the major crisis. Since about two-thirds to three-fourths of all welfare expenditures (excluding education) are allocated to the elderly and unemployment is one of the smallest programs, this seems a strange choice. Admittedly

the problems in this area have called attention to the inability of the welfare state to continue to expand, but this is not where the major cost savings are to be found.

But if one failing of the book is the narrow focus on what the crisis of the welfare state is, another failing lies in not placing the various innovations in a comparative framework. The Miller article begins this and could have been used as an introduction. But then each of the chapters that focuses on various attempts to innovate should describe which of the four categories is represented and the question should be posed as to why particular societies move in certain directions. Again, an important theoretical problem is ignored.

In sum, I found the book too descriptive and not analytical enough. At least a common framework for description should have been used.

The Power Elite and The State: How Policy Is Made in America. By G. William Domhoff. Hawthorne, N.Y.: Aldine de Gruyter, 1990. Pp. xix + 315. $47.95 (cloth); $19.95 (paper).

Jerry Lembcke
Holy Cross College

In this volume of essays Domhoff extends themes from his critique of pluralist sociology to a critique of work on the state by structuralist Marxists and state-centered theorists. His main point in this book is that, like pluralists, both of these more recent trends "abandoned the study of social power in general for more narrow concerns such as 'class structure' or 'state power' " and thereby fragmented the study of power into studies of political economy (structural Marxists) or politics (state autonomy theorists). As an antidote to the fragmentation of the field, Domhoff puts forth Michael Mann's *The Sources of Social Power* as an integrative framework within which the study of power can move forward.

Domhoff introduces the book with a brief survey of definitions and treatments of power by Hunter, Mills, and Dahl whose works form the backdrop to Domhoff's own extensive contributions to the field. Chapters 1 and 2 provide a succinct and useful review of the debates within power structure research since Domhoff's 1967 *Who Rules America?* Domhoff sorts the debates into five groups: (1) disagreements over indicators of power, (2) debates over whether leadership or structure best explains organizational behavior, (3) questions about the homogeneity/heterogeneity of the capitalist class, (4) debates over the class composition of various Western European states, and (5) how unified capitalist states actually are. In these chapters we find the origins of the acrimonious encounter

between Domhoff and his critics and a reiteration of Domhoff's denial that he is an "instrumentalist."

The book's core chapters are case studies intended to challenge Domhoff's structuralist opponents with empirical data he finds contradictory to their most important claims. He varies his tactical approach in each chapter. In chapter 3, he counterattacks directly, using studies of the Social Security Act of 1935 to show that Skocpol and other state-centered theorists are wrong in their claim that government officials were the central figures in shaping that legislation and that business leaders played a minor role. In chapter 4, Domhoff "does the Wagner Act again," critiquing his previous assessment that the passage of this pivotal piece of labor legislation rested on the support, or lack of opposition, by conservative business leaders. However, he argues that state-centered theorists are wrong to maintain that corporate elites had no role in bringing the Wagner Act into being. Domhoff finds that "the use of collective bargaining in some industries and the advocacy of it by many business leaders made collective bargaining legitimate in the eyes of Congress by 1935" (p. 66). Moreover, he finds that many of the individuals instrumental in creating the Wagner Act have been inaccurately characterized as state managers or independent experts by state-centered theorists. Some, like Lloyd Garrison and Francis Biddle, were Harvard Law School graduates with upper-class origins. Others, like University of Wisconsin economist John R. Commons, who was hired by the business-dominated National Civic Federation, were clearly compromised by their dependency on capitalists.

Domhoff concludes that to explain why the Wagner Act passed we need a theory that takes account of the segmented nature of the capitalist and working classes and comprehends the social chemistry set in motion by that segmentation. It is a response that shifts the terms of debate with state-centered theory in a way that insures further exchanges between the two camps. Less clear is the relationship between his new position and that of class-struggle theorists who have argued in regard to the Wagner Act and the Agricultural Adjustment Act that New Deal changes came from the bottom up to a greater extent than either Domhoff or state-centered theorists grant.

In other case studies Domhoff revisits his debate with Fred Block about the origins of the International Monetary Fund, examines the making of foreign policy and trade policy, and analyzes the Employment Act of 1946 to attack the idea that, through their control of investment decisions, capitalists control the parameters of government policymakers to such an extent that more direct influence in government is not necessary. The idea that capitalists rely on "underlying 'structural imperatives'" to insure their desired outcomes is, according to Domhoff, a "fairy tale."

The straightforward organization of the book around theoretical debates tested by well-defined case studies makes *The Power Elite and The State* a generally accessible book. Following the twists and turns in the history of the debates can be tedious, yet those accounts, and the inclu-

sion of empirical details around which those debates have revolved, are surely some of the most important contributions of the book. Embedded in the case-study chapters are nuggets like the author's 13-page history of the corporate liberal theoretical tradition. This is a book outfitted with a lawyerly prose appropriate to its mission. Yet in his light-hearted digs at his structuralist opponents for their brilliant self-promotion, and his teasing distinction between plain American Marxists (like himself) and the more exotic French kind, we find the rhetorical touches characteristic of Domhoff books.

Lawyers in Society. Volume 3: *Comparative Theories.* Edited by Richard L. Abel and Philip S. C. Lewis. Berkeley and Los Angeles: University of California Press, 1989. Pp. xiii + 555. $47.50.

Michael J. Powell
University of Auckland

Richard Abel and Philip Lewis, well-known scholars in the sociology of the legal profession, have edited a set of three volumes entitled *Lawyers in Society* of which this is the third. The two earlier volumes consist of detailed national studies of the history and demographic characteristics of legal professions in common law and civil law countries, respectively. The three volumes together constitute the product of an ambitious project of the Working Group for Comparative Study of Legal Professions created by the Research Committee on Sociology of Law of the International Sociological Association.

In this third volume, noted figures in legal profession studies and in the sociology of the professions draw upon the empirical analyses of the first two volumes to generate comparative theories of the legal profession. The emphasis should be on "theories" in the plural because the collection is eclectic and diverse with the authors not forced to fit their arguments into any theoretical straitjacket. The market-control model espoused strongly by Abel hovers in the background, however, with several authors either supporting or taking issue with it. The diversity of theoretical approaches contributes to the richness of this volume, although it also inevitably undermines its coherence and unity.

The book has 10 separate chapters plus a concluding essay by the editors. The topics covered range from the expansion of higher education in the 1960s and 1970s, feminist theory and the feminization of legal professions, lawyers and new class theory, the relations between the state and the legal professions, neocorporatism and professional politics, to deprofessionalization. As is common in such collections, there is considerable variation in the strength and quality of the essays. One or two really fail to make much of a theoretical contribution at all, such as Friedman's brief and rather straightforward introductory chapter and Abel's lengthy summary of the 18 empirical reports of the first two volumes utilizing

his supply-control and demand-creation model uncritically. Szelenyi and Martin's efforts to apply new class theory are somewhat strained and limited by its relevance to a brief period in legal profession history, and Neave's otherwise important piece on developments in European higher-education policy applies the concepts of deprofessionalization and reprofessionalization very loosely indeed.

The other essays are theoretically more sophisticated and more suggestive of areas for further exploration. Lewis's chapter nicely discusses the dilemmas and tensions between generalization and historical and national contingencies inevitably present in comparative and theoretical work with Lewis himself favoring a historically contingent approach, a position that is somewhat at odds, interestingly, with the deterministic market-control model relentlessly applied by his coeditor in the very next chapter. Rueschemeyer and Halliday acknowledge the historical contingent nature of state-profession relations while calling for a redirection of attention away from the market to the state, a redirection that is timely and absolutely necessary if developments in civil-law legal professions are to be taken seriously. Burrage presents a stimulating argument for the impact of revolution on the different historical paths taken by the French, American, and English professions, while Larson proffers a tentative but fertile political sociology of "lawyering" (rather than lawyers) in terms of their representation function. Moving in a different direction, Menkel-Meadow courageously, and without undue optimism, asks whether the feminization of the legal profession will also mean the feminization of legal practice and the law or simply more of the same.

Important comparative and theoretical questions, however, remain unaddressed by this collection. Despite Abel and Lewis's claim in their concluding chapter that the authors in this volume have shifted their attention from lawyering as an occupation to what lawyers do, Menkel-Meadow and Larson are in fact the only contributors who do so. This large gap may be a consequence of the sociology-of-occupations approach taken in the earlier empirical reports but unfortunately results in truncated and incomplete comparative theory. Any comparative theory of legal professions must deal systematically with differences and similarities in what lawyers do for clients and employers in different societies and economies. Ironically, Abel and Lewis themselves recognize this in their final essay and present a useful discussion of the "repertoire of lawyer functions." But it is too little and too late.

Another sizable gap is any sustained attention to legal knowledge or to lawyers' application of that knowledge, whether it be formal and academic knowledge or practical skills learned on the job. Its absence is all the more noteworthy since recent books by Freidson and Abbott have resurrected professional knowledge from the scrap heap to which it had been consigned by the monopoly theorists because of its presumed contamination with functionalism. Again, this apparent weakness may reflect the limitations of the original task and agenda set the contributors. Once more, however, Abel and Lewis's final essay seeks to correct the

imbalance with an insightful discussion of lawyers' knowledge, examining the relationship between formal and practical knowledge. If only there was more of it.

In addition to these significant gaps in coverage, there are other areas of recent critical change in the political and competitive position of legal professions, especially in Europe and North America, that deserve serious comparative and theoretical attention. One of these is the supposed deprofessionalization of law and the state of crisis that some suggest surrounds many national professions. Several of the authors do in fact touch on this question, with Szelenyi and Martin, Abel, and Larson discussing it more fully. But it is a topic crying out for sustained comparative analysis. Another such area is the convergence or lack of convergence of civil law and common law professions as a function of the apparent globalization of law and business services. The spread of multinationals and the international merger and acquisition market have encouraged the penetration of the European market for corporate legal services by large English and American firms. As a consequence, European and Anglo-American legal practice, at least at the upper end, may become much more alike.

Not that we should expect this volume to cover all the critical areas in the sociology of professions or present a comprehensive comparative theory. It represents, after all, pioneering work ably guided by Abel and Lewis and consciously limited in its scope. It is a first and necessary step. But the gaps to which I have pointed are not minor or incidental but major lacunae. Abel and Lewis's excellent final chapter admits as much. Attention earlier in the project to their clarion call in that chapter to put law back into the sociology of lawyers would perhaps have strengthened this collection. Certainly further comparative and theorizing efforts need to pay more attention to lawyering and less to lawyers.

While most of the essays in this volume have something substantial to offer sociologists of the professions, and certainly repay close attention, none really heralds a new or major breakthrough. The value of the collection lies in its practical demonstration that further advances in the sociology of the professions depend on comparative analysis of the kind and quality represented herein. Abel and Lewis, and their many collaborators, have not only pointed the way forward but led by example.

Tournament of Lawyers: The Transformation of the Big Law Firm. By Marc Galanter and Thomas Palay. Chicago: University of Chicago Press, 1991. Pp. xii + 197. $27.50.

Daniel Poor
Graduate Center—City University of New York

Corporate law is not what it used to be. Once stuffy bastions of a small group of socially connected, elite law school graduates, elite law firms

have become hierarchical, bureaucratized, entrepreneurial, large organizations. In 1960, 38 firms had 50 or more lawyers nationally (p. 22). In 1990, 21 firms had more than 400 lawyers. Many big firms have more than doubled in size since the mid-1970s.

Marc Galanter and Thomas Palay argue that the spiraling growth of the large corporate law firms since 1970 is primarily a product of the logic of the promotion-to-partnership system, adopted by firms early in the century. They see law firms as human capital sharing arrangements, where partners have a surplus of human capital in the form of reputation, client contacts, and practical knowledge and where associates possess the licensure and book knowledge to practice law, but lack clients and the knowledge gained by doing. Time constraints necessitate that partners hire associates to valorize their surplus human capital. Associates must apprentice with partners to acquire their own stock of the intangible elements of human capital that partners possess.

Yet, this sharing entails risk. In this neo-utilitarian world, agents will presumptively cheat, shirk, grab, leave, and generally act opportunistically. Following Coase and Williamson's "transaction cost" theory of firms, the authors explain the existence of law firms as a practical solution for avoiding market-exchange costs and surveillance of such a hard to value good as intellectual labor. Further, each side of the transaction provides a bond or "hostage" to bind it to the transaction. Making partner, and thus sharing in future partnership profits, is the bonus for the brightest, hardest working, nonopportunistic associates. Firms must not appear opportunistic in making partnership appointments, in order to continue attracting first-rate law graduates.

Galanter and Palay make much of the associate to partner ratio in big firms. The more associates per partner, the higher partner profits will be, other things being equal. Historically, associate to partner ratios have moved inertially upwards, from roughly 1:1 in the 1920s to 3:1 in the 1980s (pp. 59–62). Herein lies the dilemma. Partnerships must make a certain number of associates partners, to service new clients and maintain associate morale. Each new partner requires a *proportional* increase in the associate ranks—the promotion system is an exponential "growth engine" (p. 88).

Yet, "growth changes the character of the firm. Informality recedes; collegiality gives way; notions of public service and independence are marginalized; the imperative of growth collides with notions of dignified passivity in obtaining business" (p. 3). The big firms have been "transformed" by the logic of their organizational form (chap. 4). Long the pinnacle of professional self-identity, the big firms' transformation has fueled a new crisis of professionalism.

This book has much to commend it: a lucid, well-documented overview of the evolving big firm organizational form, perhaps the definitive data set on big firm growth, a clear explication of big firms' transformation and the changes in the market for legal services, and an adept use of the most valuable aspects of the new institutional economics to the

peculiar organizational form of the law firm. This alone makes the book essential reading for the sociology of law, organizations, and the professions.

It does not, however, offer an adequate explanation of large law firms. First, the normative power of the big firm organizational model must be explained. The continuing existence of low/no-growth big firms suggests that other organizational forms were available. However, the interorganizational migration of organizational models is beyond the strengths of neo-utilitarian theory. Second, it seems improbable that law students on the job market are able to acquire meaningful information about firms' promotion rates—a requirement for effective auditing. Even from within the rational action paradigm, Simon's satisficing theory draws into question the efficiency of both associates' and partners' management of informational uncertainty. If law graduates were strictly rational in firm choice, they would avoid large firms because of their low promotion rates.

Third, Galanter and Palay hint at what goes on inside these firms when real humans interact to (re)create a firm culture, yet they draw back from the very implications of invoking non-market-rational factors. They write: "the development of a firm *culture* through *social control* [which they define, quoting Parsons, as,] those processes in the social system which tend to counteract the deviant tendencies [,] and prospective monitoring will play an active role in mitigating opportunistic conduct. [Prospective monitoring identifies attorneys with] a reasonably *homogenous set* of *characteristics and motivations* . . . which [develop] through family life, schooling, training and past employment" (p. 120 and n.93; my emphasis). Social control and values are introduced deus ex machina style to make the theoretical tendency toward entropy conform with the empirical reality of organizational stability.

In short, this well-written, rigorously documented survey of large law firm growth will be usefully read by many. Its flaws come from the authors' attempts to shoehorn a complex social phenomena into rational action's partial theory of organizations.

Punishment and Modern Society: A Study in Social Theory. By David Garland. Chicago: University of Chicago Press, 1990. Pp. 312. $29.95.

Reconstructing the Criminal: Culture, Law, and Policy in England, 1830–1914. By Martin J. Wiener. New York: Cambridge University Press, 1990. Pp. ix + 390. $39.50.

John R. Sutton
University of California, Santa Barbara

To Durkheim, punishment was one of the sturdiest tentpoles of social organization—second only, perhaps, to religion. Far from being a simple

instrumental response to crime, he argued, punishment *constitutes* crime as a symbolic representation of social otherness, and binds group members together in an attitude of condemnation. But the power of Durkheim's insight was dissipated in mid-20th century sociology, and the study of crime and deviance was marginalized. Structural functionalists domesticated punishment and renamed it social control. Criminologists have treated crime as a pathology of individuals or communities, not as the normal product of normal societies. Symbolic interactionists hit upon the idea that, in their terms, sanctions amplify deviant behavior, but they shied away from theorizing links between sanctioning regimes and social organization. Crime and deviance is now a subfield of specialists, riven by conflict and unable or unwilling to speak to any larger sociological concerns.

In the last several years, however, the study of punishment has revived. For some time the strongest vein of work was Marxian in tone, inspired by Rusche and Kirchheimer's magisterially flawed *Punishment and Social Structure*. Important social histories of punishment published in the 1970s—by E. P. Thompson and his Warwick colleagues, Rothman, Ignatieff, and, of course, Foucault—lent historical depth and no small amount of theoretical confusion to the enterprise. Finally, contemporary events have given the study of punishment a sense of relevance, and even urgency. In the United States, most conspicuously, despite a long-term decline in crime rates, prisons are bursting at the seams, mainly with poor black and brown men enlisted as cannon fodder in the war on drugs. These developments, culminating in the sad celebrity of Willie Horton, call us back to Durkheim's ironic vision.

Thus David Garland's *Punishment and Modern Society* appears at an opportune time—in what we might call the adolescence of the study of punishment. As his title implies, Garland seeks to reassert punishment as a central feature of social organization. His approach is broad and synthetic, and this study can fruitfully be read as a text on theories of social control, legal change, and social organization generally. Garland rounds up all the usual suspects, devoting two chapters each to Durkheim, the Marxian tradition, and Foucault; he unexpectedly, but aptly, brings in Weber to illuminate the role of rationalization processes in the development of Western penal policies; and he moves from there into a lengthy analysis of the literature on culture, especially the foundational work of Norbert Elias.

It is here that Garland makes his most creative and daring contribution. He seeks to reinvigorate Durkheim's neglected emphasis on punishment as moral theater, but without invoking spooky notions like the *conscience collective* to account for punishment policies. Culture, he argues, is the medium that links punishment and moral order. This link works in two directions. First, following Durkheim, he suggests that punishment is a symbolic representation of wider cultural patterns. However culture is not an ethereal force, but rather a discursive project in which economic, political, and bureaucratic interests participate and

struggle for dominance. Second, punishment policies contribute to the reproduction and transformation of the patterns of moral order from which they arise. Debates over criminal justice encode conflicts over race and class, about who is deserving and undeserving of social attention, about who is in and who is out of the moral enclave.

Garland's treatment of the theorists he draws upon is erudite, faithful, and constructive. But the proof of his synthetic model is in its utility in accounting for watershed transformations in Western punishment policies: the shift from torture and execution to imprisonment as the modal punishment in the early 19th century, the bureaucratization of prisons and the rise of therapeutic ideologies in the early 20th century, and the more recent return to the punitive agenda. Garland builds on Elias to argue that the gradual rejection of sanguinary punishments and the rise of the prison are by-products of the self-conscious "civilizing" of Western society in the 18th and 19th centuries. As interpersonal violence became less common in everyday life and states became more secure in their authority, punishment was moved to the backstage of social life. The rationalization of punishment is a form of repression, in the full Freudian sense of the term: prison walls deprive us of a cathartic spectacle. Thus, argues Garland, the contemporary revival of punishment is a sort of "return of the repressed."

This book is a tightrope walk: neither an arid exercise in theoretical hermeneutics nor an original historical analysis, it is a magnificent example of *working* social theory. Yet the punchline is weak. As explanatory variables, Elias's sensibilities and Freud's tribal id are hardly more satisfying than Durkheim's collective conscience. Again I think about Willie Horton, and wonder if his unwonted fame might be explained more parsimoniously by the constituency politics of the Republican party than by some unobservable pent-up yahoo bloodlust. The turn to culture is clearly the right one; my only criticism of this fine book is that it does not go far enough. Garland ignores the literature on cultural production (names like Bourdieu, Wuthnow, DiMaggio, Griswold, and Lamont are apposite here) that treats culture as a contested strategic resource, and seeks explicitly to explain the rise and fall of dominant cultural forms. To account for shifts in punishment regimes requires more theoretical leverage than is available from interpretive approaches to culture alone.

Martin Wiener's *Reconstructing the Criminal* is not as original or powerful a book as Garland's, but it is nonetheless an important contribution to the historiography of the prison. It is intended, to use the author's words, as a "cultural history" of penal reform in 19th- and early 20th-century England. The author's use of culture seems a bit ambivalent, in ways that echo Garland. On the one hand, Wiener is concerned with the ways in which culturally embedded concepts inform both the public understanding of criminality and the approach taken by penal reformers to prisoners. The title is thus a postmodern pun: prisons, aiming to reconstruct (or reform) individual criminals, are themselves reconstructed by changing social circumstances. On the other hand, Wiener's take on

1491

culture seems a bit antique. Culture is treated more as an ambience than as a structured arena of social contest, and as the almost exclusive property of the "policymaking classes" (p. 10). But official ideology is a legitimate subject of investigation on its own, and Wiener is crystal clear about the scope of his analysis. Within that scope, this is a very fine book.

Wiener's central argument is that in the late 19th century the motivating logic of imprisonment shifted, gradually but decisively, from moralism to determinism and administrative rationality. This is not a totally new argument, but *Reconstructing the Criminal* presents it with an exceptional degree of historical nuance. Most standard accounts follow Radzinowicz in drawing a sharp line between "classical" and "positivist" criminologies: classicists like Bentham saw the criminal as rational, and punishment as a term in a cost-benefit equation; positivists saw the criminal as an irrational subject of environmental forces that could be brought under scientific control. Foucault's work implicitly erased that dividing line, treating classicism and positivism as, at most, different facets of a single power/knowledge machine. Wiener reasserts the distinction, but finds that Bentham and his ilk were primarily moralists: they sought to establish a disciplinary regime that would stimulate the innate self-governing capacities of the criminal and build character among the noncriminal public. In the late Victorian period, law was "de-moralized" in a way that we now tend to associate with utilitarianism. Faith in the inherent morality of individuals gave way to "causalism," a desperate search for external levers of control over conduct. It is at this point, Wiener argues, that the prison became a true machine, linked closely to the administrative condescension of the welfare state.

Wiener is more successful in reconstructing the prison than he is in explaining the shift from moralism to causalism. I see three factors at play in his account. First, the professionalization of policing, prison administration, social work, and domestic policymaking led to the "de-moralization" of social problems and an increasing emphasis on bureaucratic rationality. Second, the apparent success of Victorian prison policy led to demands for reform: as crime rates declined, elite attention shifted from generalized disorder to the stubborn problem of the "habitual offender," which called for a medical approach to crime. Third and most important, Wiener finds a broad cultural shift in popular views of crime and dangerousness. Crime became less romantic and more prosaic; by the 1890s, "the criminal was no longer a wicked individual but rather a product of his environment and heredity" (p. 226). Wiener documents this shift with evidence from a wide range of sources, including social policy debates, law, and middle-class literature. In one fascinating discussion, the author describes changing themes in Victorian detective novels, from Dickens's Inspector Bucket to Sherlock Holmes. I will not give away his story; it is sufficient to say that wonderful nuggets of this sort abound in this book.

Reconstructing the Criminal presents a very successful description of

the relationship between penology and culture. Hard-nosed sociologists will be less satisfied with it as an explanatory account, in large part because the author is unclear whether culture is a reflection of wider social trends or a causal agent in its own right. Most conspicuously, one misses any attention to the role of the state or any sense that prison reform was a hard-fought political issue. This was, after all, the period in which England extended the franchise to propertyless males, and, as a gesture to that constituency, began to build a comprehensive welfare state. It would only further Wiener's agenda to show how culture informed the working discourse of official actors who negotiated these transformations in electoral, social, and penal policies. But that would be a different book, and this one is fine as it is. As it turns out, that other book has already been written, by David Garland himself, under the title *Punishment and Welfare*. That study is an indispensible supplement to the two reviewed here; altogether these books attest to the renewed vigor and salience of the study of punishment and to the likely important role of culture as a source of theoretical ideas.

The State of Social Control: A Sociological Study of Concepts of State and Social Control in the Making of Democracy. By Dario Melossi. New York: St. Martin's Press, 1990. Pp. ix + 219. $45.00.

Castles of Our Conscience: Social Control and the American State, 1800–1985. By William G. Staples. New Brunswick, N.J.: Rutgers University Press, 1991. Pp. xii + 197. $37.00.

Steven Spitzer
Suffolk University

Ever since the field of social control was rediscovered and given a new and critical direction in the 1980s, attention has been increasingly focused on the role played by the state in the creation, exercise, and conceptualization of social control. Instead of assuming that social control was synonymous with the operations of an abstract and reified "society" and embedded in the routines of everyday life, the new social-control theorists turned their attentions to the highly centralized political institutions and processes that shaped social institutions, organization, and practices. Bringing the state back into the study of social control required increasing attention to the social dimensions of power and a renewed interest in the role played by conflict in the organization of society on the structural level. Weber, Marx, and other conflict theorists—once thought to be treating subjects far removed from the study of social control—were increasingly invoked as authorities in an effort to rethink the connections between the uses and abuses of power and the ordering of social life.

The two works I review here place the state at the center of efforts to describe and explain the definition, organization, and implementation of social control. While they differ greatly in their assumptions, scope, focus,

and theoretical orientation, both of these books expand our understanding of exactly how social control comes to be either embedded in, shaped by, or differentiated from the workings of contemporary states. And because they attempt to expand the dialogue about the limitations as well as the strengths of a "state-centered" approach, these works also deepen our understanding of how far we can actually go in connecting society, social control, and the state.

Dario Melossi's *The State of Social Control* is a truly remarkable work of penetrating theoretical insight. Beginning with a "reflexive" approach that recognizes the ways in which the state organizes, unifies, and rationalizes the "reality" that it purports to describe, Melossi not only demonstrates how the "concepts of state and social control belong to different intellectual traditions, and are embedded in different historical situations" but how these situations "were used for the purpose of *accomplishing* political and social order—not simply *describing* it" (p. 2).

There is no other work currently available that provides such a historically and theoretically sensitive account of how the concepts of state and social control have been invoked by political and social theorists. The first part of the book examines the genesis of conceptions of the modern state through the works of theorists such as Machiavelli, Hegel, Marx, Kelsen, Gramsci, and Freud. The far-ranging and sophisticated discussion in this part of the book illustrates how the notion of the state has always been tied to and reflective of the struggles to depoliticize the ordering qualities of the state and shift the issue of social order from the political to the social sphere.

In the second part of the book the confrontation of democracy and social control is addressed more directly as the author argues that a new form of social control is characteristic of societies in which open communication and "rational demographic composition" (Gramsci's prerequisite for democracy) come to replace censorship and direct coercion. Here we find an important clue to the shift in both the theory and practice of social control in modern states. Through an examination of the sociological contributions of Durkheim, Mead, Dewey, and C. Wright Mills, Melossi demonstrates how the concept of social control is eventually recast in the 20th century in a fashion more compatible with what comes to be understood as "democracy."

Fashioning a theory of the subject that draws upon the insights of C. Wright Mills into shifting "vocabularies of motive," Melossi offers us a way of understanding how social control and democracy come to form part of the same social process. In this sense, it is clear that social control consists far less in the direct administration of populations by a coercive state than in the creation of a discourse of what might be called "participatory administration " Here the focus is on understanding how "controlling" symbols and meanings are generated in a manner that is itself hidden from view. The problem for the type of reflexive sociology that Melossi proposes is thus not simply to describe the links between states and social regulation in democratic (i.e., not state-centered) societies; it

is also to expose and demystify those features of state-organized control that are least visible yet most powerful in contemporary social life.

While Melossi's book seeks to refine and qualify our understanding of the ways in which the state influences both social thought and social practice, William Staples's *Castles of Our Conscience* has another goal: "to offer a theoretically-guided analysis of the role of the state in the historical transformation of institutional social control" (p. 4). Instead of stressing the interpenetration of state and society, Staples attempts to isolate the actual workings of the American state within specific historical contexts. His view of the state is neither crudely instrumentalist nor functionist. He argues, rather, that "state policies reflect the actions of knowledgeable, capable human agents involved in social reproduction and (2) that such meaningful actions are always bounded by unacknowledged conditions and the unintended consequences of previous action" (pp. 13–14).

To support his thesis, Staples analyzes two major periods in the development and operation of the state in American society. Each of these periods, he argues, illustrates how social control has been influenced by changes in the form, function, and apparatus of the American state. During the first period (1800–1929), roughly designated as the age of the "liberal-capitalist" state, two stages in the evolution of state forms are identified: the accumulative state (1800–1860) and the bureaucratic state (1860–1929). The workings of the accumulative state are demonstrated through a discussion of the problem of prison labor and discipline while the bureaucratic state is depicted as operating through patterns of centralized accumulation, including the associated horrors of classification, specialization, and the absorption of locally based institutions of social control.

The second period (the era of the advanced capitalist state) is dated between 1930 and 1985 and is further divided into the period of the "crisis state" (1930–1950) and the "liberal welfare state" (1950–1985). The crisis of the 1930s to 1950s refers to the process, widely documented elsewhere, whereby institutions lost all pretense of transformation and became "warehouses for the accumulation of minds and bodies."

The contemporary period is defined, according to Staples, by the attempt to reverse the earlier accumulative pattern in favor of a strategy involving the growth of a "medical-industrial complex"—a system based on the incorporation of professional and other "private" controllers into the "welfare/health/penal sub-apparatus." In this most recent development the functions of the state are reconstituted within agencies, organizations, and practices that have traditionally stood outside the state and its spheres of influence. At this point the state is transformed in such a way that its weakening (in form) actually increases its strength (in content)—or as Melossi might put it: social control and democracy become one and the same.

The historical outline offered in *Castles of Our Conscience* is limited in important respects. The periodizations are based on a selection of

evidence that tends to overstate the similarities between widely differing patterns of development in the penal, therapeutic, and welfare systems. Similarly, the use of secondary sources from many different levels of inquiry and focus necessarily creates a loose patchwork of argument rather than a tightly researched and reasoned analysis. Nevertheless, there is a great deal to be said for the attempt of this book to document and challenge the boundaries of our inquiry into the state of social control. All theories of social control and the state must ultimately submit to the test of historical inquiry. Works such as those reviewed here make it clear that a new and far more historically sophisticated understanding of power and social order is continuing to emerge.

Crimes of the Middle Classes: White-Collar Offenders in the Federal Courts. By David Weisburd, Stanton Wheeler, Elin Waring, and Nancy Bode. New Haven, Conn.: Yale University Press, 1991. Pp. xviii + 211. $24.00.

Kip Schlegel
Indiana University

Crimes of the Middle Classes: White-Collar Offenders in the Federal Courts represents the final study from the white-collar crime project undertaken at Yale University over the past decade. Taken together with the other three volumes representing the White-Collar Crime Series, Susan Shapiro's *Wayward Capitalists,* Kenneth Mann's *Defending White-Collar Crime,* and Stanton Wheeler, Kenneth Mann, and Austin Sarat's *Sitting in Judgment,* the collection represents, in my mind, the most important contribution to white-collar crime research since Sutherland coined the term in 1939. The white-collar crime project at Yale has both added a distinct body of knowledge to the field and provided the training ground for researchers who continue to make white-collar crime their primary focus of scholarly attention.

This contribution is a fine finale. The focus of attention in this book is a description in detail of "the offenders themselves, their backgrounds and offenses, and the relation between the two: how personal situations and organizational opportunities combine to enable some offenders to do much greater damage in the number, magnitude, and duration of offenses than do others" (p. xiv). The authors give a provocative account of the movement of white-collar offenders through the legal system, including the impact of sentencing. In the final chapter the authors return to Sutherland's original discussion of white-collar crime, linking their findings to both criminological theory generally and to public policy.

The data used are from the presentence reports of a sample of white-collar offenders convicted of one of eight white-collar offenses in seven federal district courts during 1976, 1977, and 1978. The eight offenses are antitrust violations, securities fraud, mail fraud, false claims, credit

fraud, bribery, tax fraud, and bank embezzlement. Though it would have been interesting to examine white-collar crimes whose primary harm is not economic, for example, pollution or crimes of health and safety, the authors make convincing claims for limiting their offense categories. Furthermore, a valuable addition is a set of "control" offenses considered "common crimes," which include postal fraud and postal theft.

As the title suggests, the authors stress the middle-class nature of the white-collar offenses examined in this study. They note, "although they differ systematically from common-crime offenses, the white-collar crimes committed by those we studied have a mundane, common, everyday character. Their basic ingredients are lying, cheating and fraud, and for every truly complicated and rarefied offense there are many others that are simple and could be carried out by almost anyone who can read, write, and give an outward appearance of stability" (p. 171). Yet, as significant as the mundane nature of many white-collar crimes is, I found the differences among the white-collar offenses most intriguing. The data offered here call into question the tendency to lump all white-collar offenses together as a unitary phenomenon. While there are some similarities in the offenses, there are differences in terms of the seriousness of the victimization (amount of harm and spread), their organizational complexity, and in the way offenders are handled in the criminal justice system (ways that sometimes defy logic). The authors make a convincing argument that we are better off thinking in terms of a hierarchy of white-collar crimes based on organizational complexity and seriousness of victimization. Modeling opportunity structures for white-collar crime along dimensions of organizational complexity in turn adds to our understanding of the relationship between social background characteristics (status, class, occupational position) and the nature of white-collar crimes: "To oversimplify only slightly, the social background characteristics do a lot to put persons in positions that allow the conduct of organizationally complex crimes. But it is being in such a position directly in an organization or with access to it, rather than status or background itself, that is the primary mechanism enabling the commission of the offense of greatest victimization" (p. 173).

The findings presented in this volume are thorough and replete with detail. This is careful scholarship, and the authors are not prone to sweeping generalizations or careless interpretations. Rather, they examine a wide range of variables and offer myriad and thoughtful explanations for their findings. Readers will likely discover the need to reread and contemplate the information offered here. I can offer assurances, however, that on second and third reading the book continues to stimulate and agitate, in the most positive sense of those terms.

Self-Employment: A Labor Market Perspective. By Robert L. Aronson. Ithaca, N.Y.: ILR Press, 1991. Pp. xiii + 156. $26.00 (cloth); $10.95 (paper).

Marc Linder
University of Iowa

The increase in self-employment in recent years may be debatable; the growth of the literature on the issue is not. The chief merit of Robert Aronson's opuscle—its broad sampling of data on self-employment—reflects the principal weakness of the current debate: an unself-conscious, quantitative orientation unguided by conceptualization of what is being counted. By failing to open up the black box of self-employment, Aronson neglects to explain why it matters socioeconomically whether workers are employees or self-employed.

On proper heuristic grounds Aronson urges researchers to adopt the guideline that "self-employment is basically an alternative means of earning a living by the sale of one's labor" (pp. 28, xi). Yet this robust insight is submerged by misconceiving the context in which advanced capitalism subordinates seemingly independent economic agents: "Self-employment is the oldest way in which individuals offer and sell their labor in a market economy" (p. ix). This ahistorical approach obscures the critical transformation that self-employment has undergone. In particular, Aronson fails to discuss labor unions' view of the self-employed as either wage-cutting employees or embryonic employers.

Aronson is vaguely aware of the socioeconomic and conceptual problems underlying an analysis of self-employment. But his prefatory remark that "the degree of autonomy and control these workers have over their labor . . . theoretically distinguishes self-employment from wage and salary employment" (p. xii) never informs his empirical presentation. Although Aronson concedes that "the meaning and measurement of self-employment is itself something of an enigma" (p. xi), he fails to conceptualize it. Thus he reinforces the empiricist approach, which is rooted in an uncritical acceptance of the solipsistic reification of self-employment.

Unlike many empiricists, Aronson does provide a realistic account of the situation of broad strata of the self-employed by emphasizing that "the central fact about the earnings of the self-employed in the recent period of expansion is their decline from a relatively more favorable level to a relatively less favorable level in comparison with the average earnings of employees" (p. 42). He also reports the below-average income of female self-employed (pp. 60, 65–69), who have accounted for an increasing share of the self-employed in recent years. This approach leads Aronson to express caution regarding the applicability of a neoclassical framework to self-employment.

Aronson does not take a clear-cut stand on the issue of the alleged rise in self-employment. On the one hand, he notes that it is still unclear whether "the growth of self-employment represents a significant and

lasting change in the . . . behavior of labor markets"; on the other, "its reemergence is not a transitory phenomenon." Yet this renaissance of the self-employed "as a substantial component of the labor force" and "an almost spectacular revival" beginning in the early 1970s turns out to be a rise of about two percentage points as a share of all workers. The discussion is, moreover, impaired by Aronson's failure to challenge the conceptual underpinnings of the data on self-employment, which may be so infirm as to render them meaningless.

The analysis becomes particularly superficial in the chapter devoted to the impact of the state on the self-employed, which lacks a theoretical framework. Aronson's potpourri of information on the law is, moreover, studded with inaccuracies. Thus he claims that "social protection is the one area in which governments have recognized that the self-employed are subject to the same hazards as wage and salary workers" despite the fact that public insurance for illness, disease, injury, and unemployment is "virtually nonexistent" for the self-employed. Without documentation Aronson maintains that the growth of industrial homework has not compromised the employer-employee relationship: "Thus far, the federal courts have upheld the very strict standards developed under wage-hour legislation to determine the legitimacy of claims by independent contractors. Those standards . . . have been applied in . . . workers' compensation and unemployment insurance." In fact, the statutory and judicially created standards in the latter areas are much laxer.

Because Aronson views his work as a tentative effort to assemble scattered data on which others can build, it is curious that he provides no page references for any of the materials he cites—even for huge statistical works such as censuses. These omissions are exacerbated by sloppiness such as reproducing essentially the same table with the same rubrics and sources but different data (pp. 86, 108) and misstatements of the findings of articles (pp. 30, 118).

In sum, Aronson's convenient compilation of empirical indicators on the self-employed fails to broaden the cramped scope of the current debate.

Workers' Participative Schemes: The Experience of Capitalist and Plan-based Societies. By Helen A. Tsiganou. Westport, Conn.: Greenwood Press, 1991. Pp. xiii + 272. $49.95.

Trond Petersen
University of California, Berkeley

The term "worker participation" usually refers to an arrangement where workers take part in the decision-making process of the enterprise they work in and thereby influence their own situation and the outcome of that enterprise. It takes several forms, such as consultative arrangements on production processes, quality circles, employee representation on

boards of directors, codetermination in Germany, producers' coopera-
tives in England and the United States, workers' councils in Eastern
Europe, and self-management in Yugoslavia (p. xi).

Helen Tsiganou discusses worker participation in seven capitalist
countries, Norway, Sweden, England, France, Germany, the United
States, and Japan, and in six socialist countries, Yugoslavia, Poland,
Hungary, Czechoslovakia, China, and the Soviet Union.

The approach is historical and comparative. She addresses conditions
under which the plans emerge, their differences and similarities, identifies
the main initiators of the schemes, and discusses their ideological orienta-
tions (p. xii). The study has five guiding hypotheses (pp. 14–17): (1) the
labor-market conditions of a country affect the introduction and develop-
ment of worker-participation schemes; (2) increasing levels of education
and rising expectations of workers lead to demand for more interesting
work and hence worker-participation schemes; (3) certain technological
developments and economic changes might prompt changes in the inter-
nal structure of the enterprise to favor worker-participation schemes; (4)
organized labor has facilitated the development of such schemes; and (5)
the balance of power of industrial actors influences worker-participation
schemes.

The basic strategy of the book is to take the literature on worker-
participation schemes from each country, synthesize it, and then remark
on the differences between countries on the basis of readings of the litera-
ture. This is not an easy task, as Tsiganou has to spell out the most
important developments in 13 countries, tell how social scientists have
dealt with those developments, and how the developments differ among
the countries. The present book, therefore, provides no new data or
materials, it merely states and arranges what is already known. This is
useful, but if one is looking for an original piece of research on worker-
participation schemes, little will be found here.

So overall, I will say that this book is quite valuable as a synthesis of
a wide range of studies in 13 countries. It is also useful because it gives
accounts of central features of worker-participation schemes in those
countries, with each account guided by the same underlying logic of
inquiry. It is difficult to report any overall conclusions from such a book,
but the remarks about Scandinavia (Norway and Sweden) hint at the
kinds of syntheses attempted (p. 61): "What distinguishes this model of
worker participation from experiences in other countries examined in this
book is principally its extensive character which goes beyond piecemeal
changes, isolated projects, and experiments. Regardless of limitations,
problems, and barriers to workplace democratization, the Scandinavian
model has some important theoretical implications for other contexts by
illustrating mainly the importance of the extensive institutional support
which workers' participation has attracted owing to the unique social,
political, economic, and labor relations climate prevalent in these coun-
tries."

Losing Face: Status Politics in Japan. By Susan J. Pharr. Berkeley and Los Angeles: University of California Press, 1990. Pp. xv + 266. $25.00.

Hiroshi Ishida
Columbia University

Losing Face is a study about social conflicts over status issues in contemporary Japan. According to Susan Pharr, status-based conflicts arise when there are "efforts of persons of a given social status to adjust their status position vis-à-vis those above them" (p. 5). Pharr restricts her study to issues in which the statuses of the people involved in the conflict are determined by birth, that is, status differentiation by "age, sex, caste background and other attributes that are beyond the powers of the individual change" (p. 5).

The book contains three case studies representing three different status conflicts: the generational conflict between younger party members and senior members over the power structure in the conservative Liberal Democratic party, women's struggle against men over the tea pouring in a local city office, and the attempt of the former outcaste minority group (*burakumin*) to overcome discrimination in a high school. The study examines the conditions through which the status-based conflicts arose, the goals and the shape of the protesters' organization, and the reactions of the authorities to the protesters' grievances.

One of the most significant contributions of this book is that Pharr successfully unfolds a common pattern of social conflict management by Japanese authorities that can be distinguished from the Western pattern. Instead of encouraging open discourse, negotiation, and confrontation, the Japanese approach, according to Pharr, "privatizes" conflicts. Japanese authorities view conflicts as disruptive and try to avoid and contain conflicts as much as they can. If avoidance is not possible, the preferred strategy used by the authorities is to cope with conflicts case by case individually without resorting to formal channels of resolution, making it difficult for the protesters to gain legitimacy and broader support. Any concession attained by the protesters tends to be confined to that particular conflict without becoming a generalizable solution.

Pharr points out some of the socioeconomic implications of the Japanese style of conflict management. First, the privatization of social conflicts achieves political and social stability without resorting to force. The state creates neither institutionalized channels for conflict management nor general principles that may be applied to other conflict situations. By marginalizing the conflicts and limiting the applicability of a particular episode to other cases, the spread of conflict is effectively reduced, and the overall level of social conflict and the cost of major social unrest in Japan is kept low. Second, Pharr seems to suggest that Japan's economic achievement is at least in part due to the Japanese approach to conflict management. State bureaucrats were able to concentrate on the

single-minded goal of increasing economic output without being challenged by various status groups. Bureaucratic elites spent less time listening to, accommodating, and compromising with demands set forth by different groups. Pharr seems to agree with Chalmers Johnson, the author of *MITI and the Japanese Miracle,* that the relative freedom enjoyed by Japanese bureaucrats contributed to superior economic performance. The privatization of social conflict also allowed the Japanese state to have a relatively small personnel and budget size when many Western European countries expanded the size of their governments.

The general overarching tone of these concluding remarks, however, are derived from three case studies of status conflicts—conflicts based upon ascription. One would wonder whether the Japanese model of conflict management applies to other forms of conflicts, such as class struggle and antinuclear protest in Japan. The Japanese model is presented at the end of the book as if it were applicable to all conflict situations, but the book does not contain any empirical evidence related to nonstatus conflicts. Furthermore, Pharr does not explain the relationship among different kinds of conflicts. It is the author's contention that status-based struggles "cut across the lines of cleavage represented by such identities as class, religion, and region" (p. 7). However, it is not clear how the interest of a status group, for example, the women workers' interest in abolishing tea pouring, is related to the interest of the working class in general and or what the theoretical implications are of the opposing interests of different groups.

Losing Face, nonetheless, is an important book that significantly advances our understanding of the nature of conflicts over status issues by providing the Japanese example and reexamining the Western model of conflict management. It should attract a wide audience—not only of those who are interested in Japan but also of those who work in the areas of political sociology, social movement, and social inequality.

Coffee, Contention, and Change in the Making of Modern Brazil. By Mauricio A. Font. Cambridge, Mass.: Basil Blackwell, 1990. Pp. xii + 351. $49.95.

Richard Graham
University of Texas at Austin

This provocative book both expands our understanding of Brazilian history and raises important questions regarding capitalist development and the relationship of the state to society. Further elaborating the points made in his 1983 University of Michigan dissertation, Mauricio Font focuses on the political and economic history of the state of São Paulo, Brazil, during the decade of the 1920s. He takes issue on one point or another with almost all previous authors on the period. He is often in-

sightful and always stimulating, even if sometimes not entirely convincing.

By 1920 Brazil accounted for approximately two-thirds of the world production of coffee, and most of it came from the state of São Paulo. A common view is that vast plantations, worked by impoverished immigrants, provided enormous wealth and concomittant political power to their owners. The state's political elite, it has been thought, was deeply beholden to—when not actually drawn from—the planter class who also participated actively in federal politics. In 1930 a revolution dislodged that ruling oligarchy and a process of modernization began, under the aegis of industrialists, or, some would say, of an urban middle class, committed to diminishing the national economy's dependence on exports and to pushing Brazil toward capitalist industrialization, free at last from the crushing weight of the São Paulo planter hegemony.

Font insists that, far from forming a dominant and ruling class, large coffee planters were on the defensive throughout the decade of the twenties, unable to make their will felt in the government, and beset by the expanding political presence of smallholders who practiced diversified farming while also producing almost half the coffee. Export agriculture did not lead to underdevelopment or retard Brazil's progress, but rather propelled it; not, however, through aiding the accumulation of capital in the hands of planters-turned-industrialists, but through the prosperity it bestowed on yeomen farmers and their small-town suppliers. The big planters, rather than being the victims of the 1930 revolution, actually plotted its success. It is a point of view that is bound to produce controversy—indeed it already has (see *Latin American Research Review*, vol. 22, no. 3 [1987] and vol. 24, no. 3 [1989]).

Font has carried out an exhaustive survey of three major newspapers of the time. He has used them to construct a data base that allows him to quantify the number of times that coffee planters' organizations protested actions by the government or smallholders complained about their treatment. He has also used newspapers and association records qualitatively to describe points of view and positions taken. But he does not apply the canons of historical criticism to these sources, taking them all at face value. And since he has not examined similar materials for earlier periods, his claim that the conflicts he describes and the complaints of planters he details were new is doubtful, as he himself at one point half acknowledges (p. 34, n. 44). Smallholders had always predominated in new frontiers in Brazil, but he ignores the longer range to concentrate on a single decade.

At several places Font relies on dubious logic or courageous leaps of faith to advance his thesis. First, his view that the planter class lacked power principally relies on the contention that they did not always get what they wanted, namely, lower taxes, larger government subsidies for the importation of cheap labor, more effective price supports for coffee, and debt relief. But surely the measurement of their power has to be

considered in its relationship, first, to what the planters did have, and, second, to what other groups got. Nor is the fact that planters were divided among themselves as to the best policy to be pursued prima facie evidence of weakness; it may actually suggest self-confidence and strength. Also, whereas Font concentrates on identifying the social background of the opposition party, he neglects to do the same for those in power. When facts do not fit his schema, as when planters in São Paulo made alliances with their competitors in a neighboring state (p. 228), he does not question the hypothesis itself but offers vague speculations to explain away the behavior. And his implication that, if immigrant smallholders produced nearly half the coffee, they must have had half the power is historically naive and fails to ask whether they were organized into pressure groups as were planters and how much more difficult it might have been to mobilize thousands of them in contrast to a couple of hundred large planters. Sometimes Font implies that battles were fought between social groups over control of the state; at other times it seems to be an autonomous state itself against which planters contend. He presents little evidence to back his contention that planters objected to the increasingly rational-bureaucratic direction of state government; indeed, that would be especially difficult to establish since so many of those in charge of state government were also coffee planters. Finally, Font does not even consider the possibility that planters exercised a cultural hegemony over other groups and that the very categories of the political agenda might have been shaped by the planter class.

If I argue with Font about his method and his logic, it is not to deny that he has written a significant book, but rather to assert that it is a book worth taking issue with. First, he has presented much new information. Especially interesting is the demonstration of coffee planter involvement with the rebellion of 1924. He also makes a good case to show that the pre-1930 government tilted more toward industry than the conventional wisdom would acknowledge. Second, this work will prove an important stimulus to theoretical considerations of the impact of export agriculture on political development and the role of agrarian capitalists in state formation. His deliberate decision to focus on internal structures implicitly contradicts the claim of the "dependency school" of sociologists and historians. At an even wider reach, he questions whether economic dominance necessarily leads to political power. It is all the more to his credit that he raises these questions while not ignoring solid empirical evidence drawn from an often messy reality that does not fit neatly into grand theoretical schemes. This is a valuable book for both sociologists and historians.

Sociology of the Global System. By Leslie Sklair. Baltimore: Johns Hopkins University Press, 1991. Pp. xii + 269. $42.00.

Terry Boswell
Emory University

Leslie Sklair offers a theoretical framework for sociological analysis at the global level. His focus is on transnational practices (TNPs), the analysis of which requires a global perspective. The goal of his work is to "escape from the limitations of state-centrism" (p. 6) found in other theories of global, or especially, international relations. The three principal sources of TNPs are transnational corporations, especially those producing globally standardized products; a transnational capitalist class, which includes what is more commonly called the comprador bourgeoisie; and a transnational media that disseminates a culture-ideology of consumerism.

The first question I had in reading this book was, How does Sklair's concept of the "global system" differ from the more common "world-system" perspective? I was never able to answer the question to my full satisfaction. He attempts to distinguish the two by portraying the "world system as a system of nation-states and state-centrism as an explanatory framework" in contrast with "the global system as primarily a capitalist global system and the main forces in it as the transnational corporations, transnational capitalist classes and the culture-ideology of consumerism" (pp. 90–91). I was frankly baffled by this portrayal, as the workings of the capitalist world economy are the center of world-system theory. One of the more frequent and sometimes valid criticisms of world-system theorists is that they lapse into economic reductionism, understating the importance of interstate relations.

At this point, I gave up my first quest and decided that the differences in perspective were mainly "academic." Rather than pursue the distinction, I adopted a more eclectic but ultimately more satisfying approach to the book by searching for useful pieces of information and analysis. In so doing, I found numerous examples, illustrations, and discussions of transnational practices that make reading the book worthwhile.

Two discussions of TNPs were particularly illuminating. One was the discussion of the culture-ideology of consumerism, an often overlooked subject. Sklair argues that the Americanization of local media or other direct cultural imperialism, while important, is less central to generating a global culture-ideology than the promotion of consumerism itself. Regardless of content, advertising persuades people to satisfy all needs through buying goods (a point he could have made stronger by distinguishing between advertising mass consumer goods versus targeted producer goods). In addition, advertising by transnational corporations promotes a global lifestyle because the products advertised are often globally standardized. Although these products are largely American, consumerism is a global culture-ideology, rather than simply a U.S. export. Sklair

claims that transnational consumerism was particularly important in defining the revolts in the state socialist countries.

The second, and his main emphasis, concerned the question of when transnational corporate (TNC) investment leads to beneficial development as opposed to its leading to ruinous dependency. Individual case studies may point to both possible outcomes, or more often, a combination of the two. Sklair offers six variables that lead to positive development: backward and forward linkages, retained foreign currency earnings, genuine technology transfer, upgrading of personnel, upgrading of work conditions, and an equitable income distribution. Although he does not provide a complete answer, Sklair should be credited with tackling one of the most vexing problems in the analysis of dependent development. With the collapse of the Soviet model, TNC investment remains the only obvious route to development. Sklair points out that by the 1980s, "practically all Third World and most Second World countries in a position to absorb foreign direct investment are producing incentive packages to lure the TNCs" (p. 228). He concludes with the speculation that an alternative to dependent development will lie in some form of democratic feminist socialism where "the market can be created which will be efficient without being exploitative" (p. 238).

Sklair has pointed to some of the most important problems in studies of transnational corporations and economic development. This book merits reading for those interested in developing a global perspective on sociology, a perspective that will only continue to increase in relevance.

World Mortality Trends since 1870. By Jeanne G. Gobalet. New York: Garland Publishing, 1990. Pp. xvii + 140. $28.00.

George J. Stolnitz
Indiana University

This study seeks to provide "a sociological overview of theoretical and empirical work on the history of mortality" (p. xv) by presenting "a comprehensive examination of the patterns and causes of historical mortality trends" (p. xv). Empirically, its focus on socioeconomic factors affecting mortality is marked by special attention "to the effects of diffusion of ideas and the distribution of various resources within societies and worldwide" (p. 3). "Mortality research," it premises, "has heretofore suffered from haphazard, ineffective modelling and data analysis" (p. 6) in its descriptive-type preoccupations with actual and forecast trend phenomena.

The resulting volume is a mélange of several useful or suggestive contributions, insights expressed in passing that can only be assessed after much expanded analytic expansions, and numerous shortcomings of execution.

On the positive side, a review of previous research findings makes

reference to a considerable part of the main interpretative literature on international mortality trends. Interesting use is made of iterated longevity-trend comparisons between successive paired samples of more developed and less developed country (MDC and LDC) life tables. Primarily using measures of life expectancy at birth, Jeanne Gobalet compares successive pairs of derived MDC and LDC average trends for populations with life-table data available for 1870–1987, those with data for 1890–1987, and so on for subsequent 20-year intervals of starting dates extending to 1970–1987. In addition to the trends examined for overall MDC and LDC samples, analogous comparisons are provided where possible with respect to LDC regional subsamples involving Latin American, Asian, and African populations.

The major findings are of several kinds. Average longevity measures for all of the MDC and LDC sample pairings compared have been monotonically rising, regardless of the starting date used. Average longevity in the LDCs rose at a markedly more rapid rate during the third quarter of this century than was true for MDC populations. The slowing convergence of MDC and LDC average levels since the 1960s or 1970s has been associated with reduced rates of longevity gains in both regions, implying that nonlinear rather than linear statistical modeling should be applied to analyses of both sets of regional trends. Potentially, such slowing convergence points to two derivative possibilities: it could reflect ceiling-type limitations on further longevity gains in both MDC and LDC populations and persistently large differences between both sets of limits could be expected, given the unyieldingly major gaps separating their socioeconomic and developmental situations.

It is fair to say that none of these summary findings is novel or surprising; analogous conclusions have long been recognized and abundantly illustrated, though rarely if ever on the basis of the specific iterative methodology employed here.

A contribution of potentially expansive interest is the study's stress on "world-system" factors affecting national and regional longevity patterns. International diffusion of mortality-reducing mechanisms "into countries low in the world stratification system from countries higher in that system" (p. 33) can be expected from six sources: convergence of ideologies affecting social welfare norms and policy priorities; spread of economic development objectives and experiences; transmission of health technologies; international health aid programs; socioeconomic inequalities both among and within nations, and increasing strength of nation-state institutions. Hypotheses derived from these areas of emphasis are tested by means of single-equation regression models, both linear and nonlinear, using cross-sectional, country-level units of observation. The results, though mixed and erratic, nevertheless suggest that attention to extranational mortality-affecting influences could yield a needed explanatory dimension to the interpretation of MDC and especially of LDC mortality transitions.

On the study's negative side, much more adequate execution of its

stated main objectives could have enhanced its value in significant and even centrally important ways. Because it deals almost exclusively with combined-sex statistics on life expectancy at birth (with only a brief pass at infant mortality rates), the study cuts off entire major aspects of the phenomena it seeks to explain. Life expectancies at ages beyond birth, sex differentials, variable age-specific death or survival-rate patterns, the revealingly different percentage changes typical of mortality and survival statistical measures—any of these could, if explored, have enriched the study's interpretative depth and empirical relevance. Reliance on comparisons between time series of MDC and LDC average longevity measures, essentially the exclusive approach adopted, rules out a central fact, one at least equally relevant for causal analysis: contrasts found to hold between LDC life-expectancy trends since World War II and those among MDCs when these started historically from similar initial longevity levels have been far more striking than the comparisons presented. The study's repeated suggestion that declining rates of longevity changes may augur an onset of ceiling-type limitations is all too likely to be contradicted in future, given the frequency of such contradictions in the past and those in evidence at this very moment. The finding that the average regional trends derived from life tables selected at different starting dates tend to be "remarkably parallel" (p. 55) ignores an inherent biasing feature of the approach to trend construction adopted: since the life tables chosen to derive successive average longevity trends overlapped to a considerable extent, many of the same national period changes were being averaged repeatedly. Another procedure apparently followed, that of relying on visually determined rather than statistically determined trend similarities or dissimilarities, invites false or inaccurate conclusions.

Not least among the study's oversights is its failure to allow for the possibility that setbacks to developmental trends that would favor gains in longevity (such as widening education or rising income levels) may be significantly offset by expanded health policy interventions (say those involving reinforced water supply, sanitation, or vaccination programs). This issue has become especially relevant for LDCs whose development prospects have slowed appreciably or retrogressed, frequent patterns in the 1970s and 1980s.

In sum, readers of this study are likely to find it more a source of periodically suggestive, loosely treated insights than a finished research product.

Battered Women as Survivors. By Lee Ann Hoff. New York: Routledge, 1990. Pp. 289. $49.95 (cloth); $14.95 (paper).

Family Violence and the Women's Movement: The Conceptual Politics of Struggle. By Gillian A. Walker. Toronto: University of Toronto Press, 1990. Pp. 260. $40.00 (cloth); $17.95 (paper).

R. Emerson Dobash
University of Wales, College of Cardiff

During the past two decades, violence against women has been placed on the agendas of public concern and public policy. Women in most parts of the world have acted to rouse public consciousness and to respond in a concrete, material fashion to the plight of women abused by their male partners and to their valiant struggles to free themselves from this blight upon their lives. This story is at the same time one of the daily lives of millions of individuals, of the recognition of a social problem on an international scale, of the development of one of the more successful social movements in recent history, and of social change. It variously involves women who have been abused and, to a lesser extent, the men who perpetrate the abuse, along with personal friends and relatives, activists, advocates, public policymakers, social service providers, the enforcers of justice, and academic researchers; in short, the person, the community, and the state. At stake is not just the unfettered abuses of male power within the private domain of the family, but male power itself and its numerous supports within public as well as private life.

With such high stakes at risk, along with the more immediate attending gains and losses for numerous interest groups, it should not be surprising that contest and conflict have broken out, not just between the individual men and women concerned, but also between various interest groups including activists, agencies of the state, and researchers. Along with these struggles have come personal and political successes and failures that serve as inspirational or salutary lessons for those who would enter the arenas of personal, social, and political change.

Read together, *Battered Women as Survivors* and *Family Violence and the Women's Movement* provide detailed accounts of the personal and political dimensions of many of the efforts seeking change. They differ in focus. In the first, Lee Ann Hoff concentrates on the daily lives of individual women struggling to deal with the meaning, experience, and consequences of being physically beaten, intimidated, and devalued by a loved one, usually one on whom they are dependent. Gillian Walker, the author of the second, focuses on a new social movement in the process of formation and on its efforts to influence agencies of the state, particularly around the "way of knowing" about the problem (the discourse, knowledge base, competing theories, and explanations) which is viewed as a fundamental part of the process of developing specific solutions that will, in turn, shape the nature and direction of social change. These

differences in focus serve as a corrective for those who might opt for one rather than the other.

Despite their extreme differences in focus, the works have in common several features crucial to the research method used in gathering data and information. Both emphasize the importance of placing research knowledge/information/data in the context in which it exists and can therefore be interpreted and understood. Both provide highly detailed information using systematic approaches to data-gathering techniques including psychological scales, participant observation, in-depth interviews, and analysis of discourse. Both portray the dynamic development of social processes over time rather than offering the static glimpse of one point in time that so often characterizes information generated in the social sciences. They also share an awareness of the wider context in which academic knowledge, theories, and data are produced, funded, published, used, and accepted. The influences and impact of this political dimension of the production of knowledge, so blatantly obvious to those outside the spheres of the power to fund and publish, while absolutely denied by those within it, is noted by both authors, particularly in reference to the impact of the contested, counterintuitive, and counterfactual findings about the so-called equality of violence between women and men that has gained such popularity in North America and particularly in the United States.

A concern about the use of different terminology and the consequences thereof, both in describing the individuals concerned and the problem itself, is evident in both books and forms much of the detailed discussion in Walker's work. Considerable care is taken in naming the individuals as "abused women," "battered women," and "women who have been abused," and the issue as "violence against women" and "woman abuse" instead of "domestic violence," "spouse abuse," or "family violence" with its unclear referent and consequent confusion in constructing knowledge or agendas for social or personal change. It is, therefore, unclear why Walker chose to use the term "family violence" in the title of her book.

With Hoff, we enter the lives of nine battered women and 131 members of their social networks and spend one year with them. We learn how they make sense of the violence in their lives, why they stay with a violent man, and how they manage to leave. Details are provided about issues such as women's self-evaluations, comparative levels of stress, self-esteem, power relationships, victim blaming and self-blame, crisis development, and crisis management. The thesis of "learned helplessness" is not sustained. These complex insights into the woman and her predicament are then placed and understood within the wider contexts of ideologies and beliefs about women, wives, and mothers, and women's economic and material circumstances. Social networks of kin, friends, and representatives of formal agencies are scrutinized for their attitudes and actions in response to the violence. Hoff's working hypothesis is that, "a woman's social network might constitute the intermediate

structure between the subjective world of a troubled woman and the social, political, and economic institutions that uphold values and norms about women, marriage, the family and violence" (pp. 32–33). The shelter experience and the "rites of passage" toward a life without violence, rites that include homelessness, isolation, and single parenthood, are revealed in this major life transition and illustrate how personal, social, and political contexts are entwined and cannot be disaggregated if explanations and solutions are to be sought.

With Walker, we are taken through the initial stages of a new social movement, the discovery of a new social problem, and the activists' struggles with the established interests and orientations of the state and academic researchers. There is a close engagement with debates and struggles inside and outside the movement. Walker recounts the lived experience of early activists engaging each other about strategy and tactics and challenging traditional and destructive notions, particularly among state agencies and academics, about the nature of the problem and its solution. The sites for the engagement between the movement and the state are consultations and hearings held by the Canadian government. As activist cum academic, Walker moves from descriptions, action, and strategies to an analysis of the discourse used in the main government documents focusing on this issue. She highlights the divide between feminists oriented toward radical social change and those with a liberal or equal rights orientation. While the enthusiasm, dedication, and initial vision of the movement is highlighted, the overall thesis of the book is, nonetheless, one of discourse diverted and action undermined by the established interests and orientation of agencies of the state.

Warm Hearts and Cold Cash: The Intimate Dynamics of Families and Money. By Marcia Millman. New York: Free Press, 1991. Pp. xi + 191. $22.95.

Miriam M. Johnson
University of Oregon

In this entertaining and thought-provoking book, Marcia Millman contradicts Simmel's dictum that cash is cold, and shows how money is linked with the emotional hotbed forming the "deep structure" of family life. Claiming that we do not need another sentimental portrait of the family, Millman shows us how money can reveal the often unconscious and unspoken desires, resentments, and jealousies that lie beneath the surface of families who are expected to be nothing but loving. In less than 200 pages, Millman discusses both intra- and intergenerational struggles over money. She shows how wills can be used to punish and control others and how parental monetary decisions concerning children can tell siblings exactly where they stand vis-á-vis each other. She also discusses the emotional meaning of conflicts over money in mar-

riage contracts, spousal relationships, and no-fault divorce settlements.

Millman's methods are casual to say the least. She talked to an unspecified number of people, chosen in unspecified ways—friends, acquaintances, strangers, who had had interesting or unusual experiences and professionals: accountants, financial advisors, and attorneys who deal with families and money. This approach makes sense, however, for her purposes. She is not attempting to document or compare the typical behaviors of certain populations but rather to highlight the variety of situations and ways in which family members attempt to use money to satisfy their personal emotional needs and agendas.

Her generalizations and conclusions from the cases she describes are informed more by psychoanalysis than situational or microstructural analysis. In Millman's view families are made up of individuals who barely "see" each other because they look at each other through the lenses of their own emotional needs. These needs were constructed from slights, disappointments, and triumphs experienced in the families in which they grew up. As she suggests, we may think we have relegated Freud to the dustbin, but in truth "One way or another, unconscious repetition is a major element in life, whether it be repeating a formative trauma in the hope of mastering it or repeating a scenario passed from generation to generation. . . . Into these plots, money is embroidered as a major subplot representing . . . an illusion of freedom" (p. 97).

Although each case has its psychological aspect, Millman does refer to structural as well as psychological explanations for family struggles. For example, she cites Adorno's observation that the tendency of wives to belittle their husbands in private is the underside of patriarchal marriages (p. 129). And in other ways she recognizes that women are often hampered in using money to gain their ends because they have less of it than men. She concludes that women have to work at love more than men, because they do not have money.

She also deals with class differences in rules concerning the distribution and management of money in families, noting that the middle class, being the most individualistic, has the fewest rules. But these few may be changing. In an earlier generation, middle-class parents expected to give only cultural capital and a college education to their children, who were then expected to be on their own as adults. Now middle-class young (and not so young) adults are finding themselves unable to maintain a middle-class life-style without help from their parents. These changes are causing generational strain with children expecting more from their parents than their parents had expected to give.

In a final chapter Millman discusses the contradiction between the family norms of loving and equal sharing and the individualistic norms of the larger society. Here she assesses the pros and cons of prenuptial agreements and the nonpooling of money. In general, Millman recommends that family members honestly face issues about money and the emotional feelings with which it is intertwined. Precisely because money

"casts a bright light on the underlying issues in family life" (p. 174), it should not be ignored.

Because of its intimate vignettes of people's lives, its focus on those "with money" and glamorous occupations, and its psychological interpretations, Millman's book has the fascination of a soap opera and will undoubtedly attract a broad audience. For sociologists the book does have an upper-middle-class bias and, in spite of quotes from Marxists and critical theorists, does not go heavily into structural issues of any sort. However, there is nothing in the book to keep others from doing so. For me the book raised interesting questions about family solidarity in an individualistic society and about how new emotional rules about love and money might develop to meet changed conditions. Millman should be congratulated for not only entertaining but for mapping an intriguing new area for more systematic study. Finally, the book could definitely be used as supplemental reading in family classes, social psychology classes, and classes about the sociology of emotions.

Literacy in the United States: Readers and Reading since 1880. By Carl E. Kaestle with Helen Damon-Moore, Lawrence C. Stedman, Ketherine Tinsley, and William Vance Trollinger, Jr. New Haven, Conn.: Yale University Press, 1991. Pp. xxiii + 338. $35.00.

John E. Craig
University of Chicago

How well do Americans read? And what do they read? Concern about levels and uses of literacy has been broad and deep in recent years, as it has been periodically in the past. But for all the public discussion, we have known little about the dimensions and roots of the perceived crisis. This book sets out to fill this gap, and it does so in an immensely informative and stimulating way.

Kaestle and his coauthors approach their theme from two directions. First, they provide a multidimensional map of literacy in the contemporary United States and situate it historically. Moving beyond dichotomous distinctions, they develop a taxonomy of literacy levels and types well-suited to their purposes. They then draw on censuses, polls, achievement tests, research reports, and a rich array of other sources to determine how the population has fit into this taxonomy at various points over the past century. Their general conclusions are at odds with the assumptions and assertions of many contemporary doomsayers. After establishing that reading competence is above all a function of educational attainment and that median levels of schooling have risen substantially over time, they argue that there has been a substantial and more or less steady rise in literacy levels, with no evidence of decline in recent years. This does not mean that the current concern is unwarranted. Complete illiteracy is no longer a serious problem, but "functional illiteracy"

is another matter, and high-level literacy is still not widely diffused. In addition, there remain pronounced and troubling, if slowly narrowing, gaps related to race, region, and income, all of them depicted here in uncompromising terms. Much remains to be done, as the authors stress: their argument is not with the talk of a contemporary literacy crisis but rather with those who attempt to justify action on the basis of an imagined history that does not fit the evidence.

The second approach shifts the focus from the mastery of skills to the uses and potential uses of these skills. Here, to its considerable benefit, the book becomes topically wide-ranging and theoretically and methodologically eclectic. One chapter summarizes data on the consumption of books, magazines, and newspapers and its association with education (which dominates income), gender, age, race, region, and expenditures on other media and other forms of recreation. Another seeks to make sense of the reading experiences described in the autobiographies of a heterogeneous array of native-born and immigrant women who came of age in the late 19th and early 20th centuries. A third looks at mass-circulation magazines from the supply side, focusing on women's magazines and particularly on the links among publishers, advertising, gender stereotypes, and the evolution of a middle-class consumer culture. The final chapter charts the fortunes of homogenization and diversity in print culture since the 1880s, relating them to technological changes, to capitalism and its organizational offshoots, and to the shifting demographic, cultural, and political environment. This chapter concludes with a call for policies designed to shore up diversity against the pressures for standardization and homogeneity that Kaestle associates with the free market.

Although these two themes are at the core of the work, there are several subthemes and tangents that receive scrutiny. The book begins with a survey of the now sizable literatures on the history of literacy in the ancient civilizations, in medieval and early modern Europe, and in this country before the mid-19th century. There follows a splendid review of theoretical perspectives and models relevant to the analysis of reading, ranging from the overarching conceptualizations of Raymond Williams and Harold Innis to circuit models linking authors and publishers to their readers, the "issues and gratifications" approach in communications research, and the reader response tradition in literary criticism. There is a discussion of works by cultural and literary historians on American reading practices in the 19th century that highlights our relative ignorance concerning reading from the 1880s on. A short chapter offers a balanced reexamination of the test score decline of the 1960s and 1970s, concluding that the decline was more modest than commonly assumed and not attributable to changes in the character of schooling. And under the guise of an analysis of highbrow and middlebrow print culture in the 1920s, the authors introduce us to the controversial world of "readability formulas" and to two recent text-based models of readability rooted in cognitive psychology but also informed by linguistics and computer science.

The book is nothing if not wide-ranging. There is much here that will be new to anyone with interests in literacy, and there is compelling evidence throughout that literacy merits our interest. To be sure, the work pays a price for its conceptual eclecticism and its great scope. It is as much a collection of loosely linked essays as an integrated analysis, and any careful reader will spot seeming inconsistencies and loose ends or take issue with this generalization or that policy recommendation. But the flaws are minor, and the achievement is immense. It is a path-breaking or, more appropriately, land-clearing book that provides a firm and fertile base for the additional work still needed.

Novels, Novelists, and Readers: Toward a Phenomenological Sociology of Literature. By Mary F. Rogers. Albany: State University of New York Press, 1991. Pp. 324. $18.95.

Marjorie L. DeVault
Syracuse University

Novels, Novelists, and Readers explores the experiences of those involved in making and reading novels and maps out research sites that would constitute a phenomenological sociology of the novel. Using a "social worlds" perspective, Mary Rogers addresses topics such as the work of novel writing, including the relationship of novelists' experiences to the texts they produce; reading, especially in relation to readers' everyday lives; and the commercial world of publishing. Sensitive to the pitfalls of positing a "world" that appears overly unified and coherent, she makes it clear in her analyses that she thinks of writers and readers as participants in multiple, overlapping worlds. This use of the social worlds perspective allows for fluid analyses that point to persistent, but not overly determinative, structures organizing writers' and readers' activities. Rogers concludes with some discussion of parallels and differences between novel making and sociological representation, as a move toward a more general analysis of human processes of making meanings.

Rogers makes an intriguing phenomenologically based contribution to the sociology of literature by taking the imaginative activity of novelists as amenable to sustained investigation. While she does not ignore the institutional and economic structures that are the focus of many sociological analyses of literary work, she treats these more commonly studied aspects of literary worlds as a context for the activities of observing, experiencing, imagining, and then writing or reading, and examines the social character of these activities that are usually taken to be private and idiosyncratic. A particular strength of the book lies in Rogers's sustained treatment of gender and her (somewhat less consistent) treatment of racial/ethnic divisions as constitutive elements of social worlds.

In keeping with the ambitious scope of this project, Rogers draws on an impressively—and sometimes frustratingly—wide range of data

sources. At the center of the analysis is a "sample" of novels that she constructed through a multifaceted selection process and undertook to read (or, one presumes, reread) over a period of several years. She casts her net broadly, including formulaic fiction—romances, mysteries, and westerns—as well as best-selling novels from the past one hundred years and works that are by various measures part of a "literary heritage." From her readings of this sample, Rogers constructs a rough typology of novels and sketches some possible (though for me, often unconvincing) generalizations about the relations between types of novels and their connections to everyday life. Other sources for the larger analysis include biographies of novelists and authors' own discussions of their craft as well as the writings of publishers, book reviewers, and literary critics. These other sources often enliven the text, but they work somewhat ambiguously. Rogers does not explain how she selected them, and she uses them inconsistently, taking them sometimes as the claims of social world participants and sometimes much less critically. (She is careful to note, for example, that the stance toward realism adopted by male authors during the 19th century should be read, at least in part, as useful rhetoric in a gendered move to claim the high ground of novel writing for men; she takes largely at face value, however, the kinds of comments on suffering that can be found in many literary life stories.)

Although this book is based primarily on brief discussions of entire novels (or entire bodies of work by novelists or even types of novelists), I was most interested in the analyses based on closer work with particular texts. For example, Rogers develops a thoughtful account of the social character of novel reading: drawing from reader-response criticism and Erving Goffman's account of strategic interaction, she highlights distinctive features of interaction through a textual medium, exploring such issues as the textual resources available to novelists and the strategies they use within their texts to educate readers about the genre.

The concluding discussion of parallels between the worlds of literature and social science seems somewhat at odds with Rogers's more grounded analyses of novelistic worlds. The attempt to embrace both kinds of worlds pulls her toward such broadly sketched characterizations that the only possible claims are relatively unsatisfying. The power of storytelling, perhaps, pulls Rogers (and many other sociologists) toward novels. But the impulses of sociology pull away from the particularities of storytelling. This central tension, I think, constructs the greatest challenge in the kind of project Rogers has envisioned and perhaps helps to explain why she has been only partially successful.

The Myth of the Madding Crowd. By Clark McPhail. Hawthorne, N.Y.: Aldine de Gruyter, 1991. Pp. xxx + 265. $44.95 (cloth); $21.95 (paper).

Pamela E. Oliver
University of Wisconsin

The main point of *The Myth of the Madding Crowd* is easy to state. Past theorists have assumed that the thing to be explained is why people in a crowd act unanimously. They have debated whether the unanimity was caused by the temporary creation of a group mind that made people suspend their normal rational thought processes or whether it was the product of the convergence of like-minded people. Combining the few prior observations of actual crowds with his own hundreds of hours of observations, Clark McPhail argues that unanimous action is exceedingly *rare* in crowds, and thus that theories of crowds centered on explaining unanimity are explaining a phenomenon that does not exist. He also cites clear evidence of planning and coordination that gives the lie to ideas of crowd "mindlessness." He accuses almost everybody of "armchair theorizing" without reference to the data.

This thesis could be developed as a slash and burn on the previous literature, but it is not. Chapters 1–5 review the work of LeBon, Park, Blumer, Allport, Miller and Dollard, Sherif, Turner and Killian, Couch, Berk, Tilly, and Lofland. Each theorist is treated carefully and thoughtfully, as located in a particular historical and intellectual context, and with appreciation for the insights and contributions, details, nuances, ambiguities, and self-contradictions of each argument. There are many clear figures showing the structures of arguments, and extended quotations from and citations to the empirical literature that provide a basis for understanding both the instances theorists were trying to explain and McPhail's critiques of them. These five chapters are a highly satisfying read about the history of sociology. They tell us where we are coming from, identify the enduring questions and issues that underlie theories and debates, and set up a reformulation of the problem.

In the last two chapters, McPhail reconstructs the problem. In chapter 5, he reviews and rejects extant definitions of crowd and collective behavior as too theory-laden and misleading. He replaces crowd with the neutral and more specific term "temporary gathering." He redefines collective behavior neutrally as "two or more persons engaged in one or more behaviors judged common or concerted on one or more dimensions" (p. 159, punctuation altered). He distinguishes behavior-in-common (doing the same thing) from behavior-in-concert (doing different things together). He inventories 40 specific behaviors that have been observed in temporary gatherings (e.g., locomotion, gesticulation, conversation). He develops some language for distinguishing among types of gatherings and for aggregations of gatherings. The chapter contains many vivid examples of all these terms and concepts.

In chapter 6, McPhail builds on Mead and Powers and develops a "sociocybernetic" model for individual and collective behavior. In it, people use reference signals to control their behavior so as to respond to or coordinate with the behavior of others. The model explains what really happens in temporary gatherings. It tells how people coordinate a chant, whether it is "burn, baby, burn" or "go, team, go." It tells how a march works, be it religious or political. Empirically, the argument is that people are basically purposive and figure out how to coordinate their behavior with others to accomplish their ends. It is not that people are cost/benefit maximizers or that they are emotionless, but that coordinated behavior occurs through attending either to a common source or to each other. The models in this chapter are simple and do not include conversations to coordinate complex differentiated action, like capturing, torturing, and hanging a black man or breaking into and looting certain stores and not others during a riot, but McPhail's writing throughout the book has made it clear that these would belong in his long-term agenda.

The behavioral model applies equally to football games and political demonstrations. If read only superficially, this model may alienate some scholars who work on social movements, because it seems so mechanical and apolitical. Passion and courage are not there, nor are meanings and reasons. People are purposive, but we do not know their purpose. McPhail is compelling when he argues (implicitly) that, at the microlevel of action and the coordination of action, the larger political or social context makes no difference, and that we should give up looking for something distinctively different about "collective behavior," or about politically significant behavior at the behavioral level.

But this does not make the larger context irrelevant. We know that social-change goals, emotions, political opportunities, discursive frames, collective understandings, and a sense of history all matter. What we need to do is to look for the ways in which these "big" effects affect the microprocesses. We will find many of these outside McPhail's models. The conversations and agreements that permit coordinated action within a temporary gathering often occur "offstage" to that gathering, at a prior time and another place. Also offstage are the informal conversations and mass communications that create the shared meanings and understandings that lead people to attend to assembly instructions, form a temporary gathering, and orient their reference signals to a common source.

To be sure, the models in this book are not the whole story. But they are a very important part of how things work, and our subsequent work will be much better if we take that understanding and build on it.

Modernity and Ambivalence. By Zygmunt Bauman. Ithaca, N.Y.: Cornell University Press, 1991. Pp. 285. $43.50.

Eli Zaretsky
University of Missouri at Columbia

According to Zygmunt Bauman, "modernity," the cultural and social project of the Enlightenment, is characterized by its drive toward order, design, management, naming, and segregating. Building on his argument in *Modernity and the Holocaust* (Cambridge: Polity Press, 1989) Bauman identifies its prime agent with the "gardening-breeding-surgical" ambitions of the nation-state. Communism he deems the most literal exponent of modernity, a "counterculture" seeking to realize rationalizing visions that in capitalist states remained only half begun. Bauman's originality lies in defining the counterpart to this project not as chaos or disorder but as ambivalence, a sphere of social action characterized by polysemy, foundationlessness, cognitive dissonance, and contingency.

Bauman's major example of modernity is the assimilating tolerance of liberalism. In his view, tolerance toward (assimilating) individuals was always linked to intolerance aimed at collectivities, above all at their value-legitimating powers. Modern culture, therefore, characteristically depends on the distinction between inside (ordered, accepted) and outside (disordered, unacceptable). The nation-state, Bauman argues, "is designed primarily to deal with the problem of strangers not enemies" (p. 63). Drawing upon Nicholas Luhman's argument that "with the adoption of functional differentiation individual persons can no longer be firmly located in one single subsystem of society, but rather must be regarded a priori as socially displaced" (p. 201), Bauman identifies "strangerhood" as the universal condition of modern individuality and personal life.

Ambivalence, then, is not the condition of the outsider but of "the stranger" who may or may not gain acceptance. The drive for order and inclusion generates ambivalence and contingency as surely as industrial development generates waste. The crucial psychological and social terrain carved out by Bauman is that of "undecidability," experiences and institutions which, like Derrida's *pharmakon,* meditate against the binary opposition, either/or. Postmodernity, in Bauman's analysis, is therefore modernity aware of and accepting of its inevitable underside.

Bauman's analysis centers on the case of German Jewry. He uses this example to argue that the neurotic, dissimulatory structure of ambivalence is that of the double bind: the closer German Jews came to assimilation, the more their Jewishness came into prominence. Living on the border between either and or condemned them to continual self-scrutiny, to "life-long and never conclusive examination." *Self*-scrutiny articulated with the elitist structure of the Jewish community; a psychology of shame directed at the self complemented a sociology of embarrassment

and digust directed at "others." Given the oxymoronic logic of assimilation, no cause for shame could ever be isolated. As the writer Jacob Wasserman put it, "others enjoyed a credit account . . . I, however, had to present my credentials every time, to stake my whole fortune."

Bauman claims that the universal experience of ambivalence and contingency, sustained neither by indoctrination nor by a reign of experts, renders modern societies "spectacularly immune to systemic critique and radical social dissent" (p. 261). In the absence of such critique, a philosopher like Derrida is important because he "restores the indeterminate to its rightful status of the ground of all being" (p. 189).

I have several objections to Bauman's brilliantly written and important argument. First, Bauman defines the state in terms of its ordering or rationality function but does not address the relation of the state to the market. While the market certainly gives rise to ordering and rationalizing tendencies (e.g., centralization, planning) it also generates counter-normalizing tendencies—both those that are irrational in a Marxian sense and those that counter the hegemonic logic of the state.

In addition, Bauman too easily equates the state with the nation. Even if we accept his Foucaultian identification of the state with its normalizing function, we lack any explanation of why national groups seek to maintain a distinction between inside and outside, an explanation such as Freud offered in his *Group Psychology* and *Civilization and Its Discontents*. Even in Bauman's example—Germany—the state through part of its history protected collective minorities (not just individuals) against the nation. Both of these objections converge if we take not Germany but the United States as our case study for "modernity." In the United States, on one hand, the market with all its disorder has the greatest possible autonomy vis-á-vis the state and, on the other, there is no racially based or homogenous nation or people capable of upholding the inside/outside distinction for long.

Another objection concerns the use of the Jews as a case study. Bauman notes that a considerable part of the sociological theory of modern assimilation has been based on the Jewish experience, but his own main reason for choosing the Jews—the great number of self-conscious accounts they produced—is not adequate. In my view, the reasons for the particular salience of the Jewish experience in theories of 19th- and 20th-century modernization and assimilation have not yet been articulated.

Finally, Bauman's one-sided characterization of modernity necessarily leads him to an uncritical, indeed celebratory, account of postmodernity. The Enlightenment was a more complex and internally divided phenomenon than Bauman makes it out to be. Its most distinctive characteristic was, arguably, not its ordering and rationalizing project (which characterized many premodern empires as well) but rather the new emphasis on *self-criticism,* on the examination of one's own assumptions whether by communities (science, philosophy, art) or individuals. This quality, which distinguished modernity from the first, is the necessary basis for

the tolerance of difference which, for Bauman, only entered Western thought in the 1970s.

Modernity and the Holocaust. By Zygmunt Bauman. Ithaca, N.Y.: Cornell University Press, 1989. Pp. xiv + 224. $29.95.

Moishe Postone
University of Chicago

Zygmunt Bauman has written an important and thought-provoking book on the Holocaust and modernity. He argues that the Holocaust must be understood as a central event of modern history and not as an exceptional episode that represented a historical regression to barbarism. It was based upon the technological and organizational achievements of an industrial bureaucratic society: modern processes of rationalization, which substitute organizational discipline for moral responsibility, and instrumental rationality, which is one of means rather than of ends (pp. 13–26). The Holocaust was a possibility rooted in essential aspects of modernity itself and, as such, reveals its destructive potential and negative moral possibilities (pp. 26–29).

Such an interpretation requires reconsidering analyses of modernity as a civilizing process (Elias), or as a process of the progressive rationalization of all spheres of social life (Weber). This, in turn, requires rethinking sociology, the theory of modernity, itself. Bauman claims that sociology has not adequately confronted the challenges raised by the Holocaust, in part because sociology participates in the same scientific culture of modernity, shares its emphasis on technique, its propensity for social engineering, and its understanding of rational action (pp. 4, 29). Relatedly, sociology has bracketed moral considerations by endorsing value-freedom and means-ends rationality, and by presupposing that morality is socially grounded. Such an approach, with its assumptions of progress and of socially based morality, will tend to view Nazism as an anomaly, a result of social breakdown. Bauman contravenes this view and calls for a reconsideration of modernity and of the nature of sociological thought.

Bauman begins by arguing that the Holocaust cannot be explained in terms of anti-Semitism alone. He distinguishes racism and anti-Semitism from "heterophobia" and prejudice, and notes that anti-Semitism characteristically is directed against a group it considers anomalous, outside of the "normal" classification systems of society, culture, and religion (pp. 32–45, 81). In the modern period, anti-Semitism became racist, and the Jews became a central focus of antimodernist energy (pp. 46–60). Racism, for Bauman, is a modern form used in the service of nonmodern struggles (p. 62). Racism itself is tied to modern conceptions of social engineering, to the idea of creating an artificial social order by changing the present one and eliminating those elements that cannot be changed as desired (p. 65). Exterminatory anti-Semitism was eminently modern,

according to Bauman. It was dependent on the modern phenomena of racist theory and the medical-therapeutic syndrome (p. 76). Moreover, it needed a modern means for its implementation: modern bureaucracy.

Bauman then analyzes the Holocaust in terms of modern techno-logical-bureaucratic patterns of action and the mentality they generate and reproduce (pp. 95 ff.). His insightful and compelling analysis focuses on the greatly increased efficiency of action that results from the mobiliz-ing and coordinating potential of bureaucratic administration and on the effects of bureaucracy on individual morality. Analyzing the latter in light of the Milgram and Zimbrado experiments, Bauman argues that the rise of instrumental rationality and bureaucratic organization allowed for a new form of crime by dimming the causal link between the actors' actions and the victims' suffering and by providing a substitute morality of "duty" and "discipline." Historically, these forms of rationality and organization also promoted cooperation by the Nazis' victims.

As opposed to what he considers the dominant sociological position (that society is a moral institution that humanizes individuals), Bauman argues that inhumanity is rooted in social relationships. As the latter are rationalized, the possibility for the social production of inhumanity increases (p. 154). Bauman concludes by calling for a new sort of sociolog-ical theory that incorporates, rather than excludes, ethical considerations. He asserts that morality is not a product of society, but is rooted in the human condition of "being with others" (Levinas) and is manipulated by society.

Bauman's arguments are important and powerful, as is his demand that sociological thought respond seriously to the Holocaust. Neverthe-less, several aspects of his book are problematic. Bauman argues that the Holocaust cannot be understood in terms of anti-Semitism, and yet implicitly indicates that biologistic, exterminatory anti-Semitism was a necessary condition of the Holocaust. He does distinguish anti-Semitism from prejudice and "heterophobia" in general, but does not sufficiently distinguish modern anti-Semitism, which culminated in extermination, from other forms of racism, which do not implicitly point toward the annihilation of the other. Had he done so, Bauman could have more successfully overcome the opposition between "functionalist" positions, that argue that the Holocaust emerged contingently as a goal in the course of Nazi rule, and "intentionalist" positions, that argue that the Holocaust was planned from the beginning (pp. 15–16). The real histori-cal problem is not simply whether the Nazis had always intended to exterminate the Jews, but that a program of complete extermination had become a conceivable possibility. On this level of consideration, a closer relation exists between modern anti-Semitism and the Holocaust than Bauman's account indicates.

Additionally, Bauman's discussion of the relation between modernity and anti-Semitism is not fully satisfactory. He characterizes modern anti-Semitism as the expression of antimodern phobias in modern form

(pp. 46, 62). While this widespread interpretation may be valid for Poland, for example, it is less convincing as a characterization of Nazi anti-Semitism, which only rejected some elements of modernity and emphatically affirmed others (technology, industrial capital, etc.). Indeed, the difficulties in adequately characterizing Nazi anti-Semitism as antimodern reveal the limits of "modernity" as an analytic concept. Bauman is critical of the way in which sociology has treated that concept. Yet, by characterizing anti-Semitism as essentially antimodern, he implicitly adopts the position toward modernity that he criticizes.

One aspect of Bauman's book is very puzzling. He argues that instrumental rationality and bureaucratic-technocratic domination are hallmarks of modernity, that structural similarities exist between those characteristics of modernity and sociological modes of thinking, and that the Holocaust revealed the negative dimensions of precisely those aspects of modernity. Similar arguments were also developed by Adorno, Horkheimer, Marcuse, and others associated with the Frankfurt school. Yet, surprisingly, Bauman does not comment on the evident parallels between his position and theirs. He does mention Adorno's contribution to *The Authoritarian Personality,* but misinterprets it as an attempt to reduce Nazism to the abnormal psychological characteristics of the people who tended to become Nazis (pp. 152–53). In fact, Adorno's larger argument was that the bureaucratization and rationalization of modern capitalist societies in the late 19th and 20th centuries had also transformed the structures of authority and the family and, thereby, had generated historical changes in character structure. These changes included a shift in superego functions, similar to that noted by Milgram, from evaluating the goodness or badness of actions to assessing how well or poorly one is functioning in an authority system (p. 100).

Although Bauman's position parallels central aspects of that of the Frankfurt school, they differ fundamentally as regards the issue of society and morality. Bauman does not appropriate their sophisticated attempts to consistently regard humans as culturally, socially, and historically constituted by socially grounding both conformist and oppositional worldviews. Such an approach would have served his intentions well. Instead, Bauman adopts Durkheim's opposition of individual and society as one between morality and amorality and simply reverses its terms. The individual now becomes the "pristine source" of morality, and society is regarded as that which undermines it (p. 199). This attempt to recover the possibility of moral evaluation has the weaknesses of all conceptions that posit a human essence outside of determinate societies and cultures and treat the latter as extrinsic to individuals.

In spite of these problematic aspects of Bauman's book, it is a forceful and convincing argument that the Holocaust requires a critical reconceptualization of modernity and of social theory and, as such, represents an important contribution to contemporary sociological thought.

How Does Social Science Work? Reflections on Practice. By Paul Dies-
ing. Pittsburgh: University of Pittsburgh Press, 1991. Pp. xii+414.
$29.95.

Lawrence E. Hazelrigg
Florida State University

Paul Diesing offers a propaedeutic of, and incentive for undertaking,
greater self-consciousness of reflexivity in the practices of social science.
The instruction is bracing, both for its wit and for the persistence of its
mirror, which is turned inward often enough to make clear that this is
not a dyspeptic exercise in truthmongering. A genuine concern for the
health of social science could not easily be stated more profoundly than
in this book.

Diesing builds his instruction around three questions: What are "the
actual goals of the various current research methods," and how do they
compare with formal professions of aim? "What social, cognitive, and
personality processes occur or should occur during research, and how
do they contribute to the outcome?" "What persistent weaknesses and
dangers appear in research, and what can we do about them?" (p. ix).
The three questions and their intersections are continually in attention,
although each is the primary focus in one of the book's three parts.

The five chapters of part 1 address different ways of formulating an
answer to the first question. Diesing has achieved a nicely articulated
overview of most of the "philosophy of science" positions staked out and
debated since the 1920s—e.g., logical empiricism, Popper's arguments
and reactions to them, Kuhn's communitarianism, Stegmüller's structur-
alist analytics, the pragmatisms of Dewey, Kaplan, and Churchman, and
some strands of hermeneutics. Written in and for the idioms of social
science rather than those of academic philosophy, the chapters strike a
good balance between evenhanded description of the given proponent's
account of the proper goals and actions of science and translations of
those accounts into actual practices of social science. Diesing makes
especially effective use of Popper's inconsistencies, Feyerabend's "anar-
chistic" pluralism, and hermeneutic insights into strategies or quasi-
strategies of (mis-)interpreting the theories of data constructions put for-
ward by scholars outside of one's own research community or theoretical
perspective.

The real excitement of Diesing's book comes in part 2, which focuses
primarily on the second question (processes of doing research) and from
there takes up each of the others. If the five chapters of part 1 may be
read as a record of hopes and promises of a "pure reason" proclaimed,
puzzled over, and then successively dashed, the five chapters of part 2
are a report of a "practical reason" taking stock of itself, sometimes
wittingly, sometimes not, and frequently under the impression that it is
"pure reason" speaking. Diesing makes the expected tour of social-
science studies of science—e.g., the Mertonian program, some of the

more externalist perspectives, recent ethnographic and constructivist studies (e.g., Bourdieu, Latour). But the greater interest of part 2 comes from Diesing's deft weaving into this "science of science" fabric a variety of insights garnered from studies (cognitive science, AI research) of how people with differing dispositions actually apply skills of reasoning in different situations, and of the extent to which those skillful applications depart from tenets of formal logic and standard professions of scientific method. Diesing marshalls considerable evidence to the effect that those departures (e.g., availability bias, anchoring bias, fundamental attribution error) are hardly smaller among scientists than among other members of the population. His point is not that discriminations in quality of research cannot be made, nor that improvements of practice are either impossible or bootless. To the contrary, he argues that both are possible if we, as sociologists, derive useful lessons from this evidence. (This saves him from playing an inconsistent game—e.g., using social-science evidence that most people have "low reasoning power," by standards of formal logic, as evidence that social-science research tends to be informed by low reasoning power.) All in all, however, the derived lessons "should teach us humility; but Merton's norm of humility will not correct over-confidence, since confidence is a matter of personality dynamics. And since people who are successful in the social science rat race tend to have a great deal of self-esteem, the outlook for improved objective knowledge is poor indeed" (pp. 271–72).

Part 3, "Putting It All Together," is somewhat disappointing in value-added terms, although this is partly because the earlier chapters often telegraphed ahead. Nevertheless, here as elsewhere, Diesing pulls into a coherent picture insights from a variety of exemplary cases of research practice, drawn seemingly at will from many disciplines. What does the picture say? Ignoring details and nuances, it says that the practices of social science are not much different from other occupationally organized task-oriented practices. Social science's product, multiple and contradictory truths, "live in the practices and understandings of a research community, not in particular laws, and when that community peters out," so, too, do its truths (p. 364).

If that message irritates you, I urge you to read Diesing's book. (The evidence on selection bias says you probably will not, however.) If the message is confirmatory, read the book anyway; its assembly of materials is unlike any other I've seen.

Social Theory for Action: How Individuals and Organizations Learn to Change. By William Foote Whyte. Newbury Park, Calif.: Sage Publications, 1991. Pp. vii+301. $36.00.

Burke D. Grandjean
University of Wyoming

Few scholars can tell a tale the way William Foote Whyte can, making the subtleties of social life so comprehensible. Fewer still ever pull together the common themes from a lifetime of research, as Whyte has done here, seeking insights for sociological theory and practice.

In the first half of the book, Whyte recounts some of his best tales from decades of work in agricultural development and industrial relations. In both contexts, he has repeatedly found that participation in decision making by those of low social rank contributes to attaining collective goals that could not be reached with more traditional structures of authority.

In Guatemala, for example, trials of new agricultural methods were moved from the ideal conditions of centralized experiment stations and into the rocky hillside fields of working farms. With farmers participating in the design and execution of the experiments, results were directly applicable to their needs, and successful innovations diffused almost at once to neighboring farms. At the Xerox Corporation, management proposed to economize by substituting outside purchase of some components for in-house production. Instead, teams of workers devised plans for reorganizing work that reduced costs up to 40%, saving hundreds of jobs. In the Basque region of Spain, the Mondragón network of worker cooperatives maintained high employment during recession, then resumed an impressive growth in sales. And in a variety of research projects, Whyte and others enlisted subjects of their research as active participants. Benefits for both the scientific value of the research and its practical application are arguments for the method that Whyte calls participatory action research.

Such tales may not sound new if one has been able to keep up with Whyte's voluminous writings during the past half-century. Over one-sixth of the entries in the book's bibliography are his own works. Still, like most good tales, these do not suffer from the retelling. Here they are juxtaposed in ways that highlight their common themes.

Nor is the effort at theoretical synthesis entirely new. Whyte says frankly that part of the motivation for the first draft of the book was "the urge to get into print once again some theoretical ideas that I thought had not received the attention they deserved from my colleagues when I had published them earlier" (p. 271). Now refined and sharpened, these theoretical ideas also bear the retelling well.

The second half of the book aims to explain why and how participation works to achieve collective goals. The account of *why* it works is based on Whyte's theory of transactional relationships (akin to exchange theory) that is summarized in several chapters. Collective action often has multi-

ple objectives rather than a single goal. If group action achieves multiple goals, it can have joint pay offs that reward actors at various social levels. When low-level actors participate in setting some objectives, their motivation is increased by the prospect of attaining rewards they value. Participation also enhances productivity, at least if productivity is defined to reflect the reduced supervisory costs when workers are self-supervising.

Explaining *how* participation becomes established involves the notion of organizational learning. This, for Whyte, is the conversion of social inventions (new, often temporary realignments in social forms) into more permanent social structures. Whyte provides many examples of organizational learning. Yet his theoretical explanation of how and when such learning occurs, or fails to occur, is less fully developed than his treatment of why participation works.

Most generally, the book makes a case for bringing applied sociology into the intellectual mainstream of the discipline. Of course, a greater concern for practical applications might increase the social usefulness of sociological research and enhance the employment prospects of sociologists. But Whyte's point is that applied research also makes essential scientific contributions to the development of sociology, grounding theory and data more firmly in the social realities of the subject being studied. That point could hardly have a more persuasive proponent. Even the parts of the argument that could use further refinement still serve as a promise that the present work will not be Whyte's final word on the matter.

Taking Society's Measure: A Personal History of Survey Research. By Herbert H. Hyman. Edited and with an introduction by Hubert J. O'Gorman with the assistance of Eleanor Singer. New York: Russell Sage Foundation, 1991. Pp. xxiv + 257. $34.95.

Michael Pollak
Centre National de la Recherche Scientifique

In his introduction Hubert O'Gorman reminds us that this last book by Herbert H. Hyman has remained unfinished. Although Hyman was able to revise the published chapters several times, important chapters of the planned book are lacking. They were to have analyzed the diffusion of survey research outside the United States. Only Hyman's photograph on the cover page, taken in Cairo during a 1965 UN mission recalls this crucial chapter that somebody else will have to write.

In part 1, Hyman describes the war years; in part 2, the postwar years. The major merit of the book is its demonstration of the crucial importance of the war years not only for natural sciences, but also for social sciences in general, and large-scale empirical research in particular. Two themes were of importance to the commissioning agencies: the reac-

tions of civilians to the war effort, deprivations, and separations and the reactions of soldiers in combat. Among the most interesting findings Hyman insists on is how methodological reflections can advance theoretical interpretations, for example, such effects as the color of the interviewer in research on political attitudes or the implications of an ethnic mix in combat units for research on stereotypes and attitude change.

A most touching chapter is the one on "The Bombing Surveys: Germany and Japan." What were the effects of intensive bombing on the civilian population? Did it shorten the war effort, did it accelerate surrender? Clearly the surveys show similar human reactions in two extremely different cultures. These surveys based on open-ended interviews are an important contribution to the study of behavior in extreme situations. During the war in Vietnam, Hyman tried to "remark on the immorality of bombing civilians when one knew it served no strategic purpose." He did this without any success: an illustrative example of the limited power of knowledge running counter to established conceptions of the people in power.

The last two chapters concentrate on the institution building of the National Opinion Research Center and the Bureau of Applied Social Research. In this history Hyman's personal trajectory from psychologist to methodologist, and later on to sociologist, is a typical career pattern that has shaped a generation of passionate pioneers in the field. But this interdisciplinarity irritated traditional academics. This successful field for practical applications has never really been accepted by the university establishments. Its fight for academic recognition was a difficult undertaking with limited success.

The Origins of American Social Science. By Dorothy Ross. New York: Cambridge University Press, 1991. Pp. xxii + 508. $29.95.

Adam M. Weisberger
Colby College

The Origins of American Social Science by Dorothy Ross is a comprehensive and penetrating examination of the emergence of the academic disciplines of economics, sociology, and political science in the United States. Her analysis, spanning the period from Reconstruction to the Great Depression, brings to light the fundamental ideological assumptions shaping the intellectual traditions of American social science. Buttressed by massive amounts of primary and secondary materials, her investigation situates American social-scientific discourse within its cultural, political, and institutional contexts and also reviews the works of a series of exemplary figures.

Ross's central argument is that the distinguishing characteristics of American social science are deeply bound up with a nationalist ideology of American exceptionalism, the belief that America occupies a special

place in world history. Ross's aim is to historicize and debunk this ideology. "While claiming to describe the American world as it was, exceptionalism instead distorted that world, providing a simplistic and idealized vision of the United States and exaggerating American uniqueness" (p. xviii). American exceptionalism, claims Ross, is premised on beliefs originating in Puritan, liberal, and republican traditions that America's republican institutions and economic opportunity had spared America the European maladies of socialism and class conflict and "set America on a millenial course" (p. xiv). American social scientists, committed to individualistic and liberal market values denying reality to class inequality, supplied intellectual legitimation to a capitalist political economy by seeking to root it in unchanging laws of nature or history. Exceptionalist ideology assumed a hegemonic position in American social-science discourse, containing ambivalence toward capitalist development and conflicting Whig and Jacksonian political traditions. The dominance of exceptionalism marginalized the alternative visions and dissenting voices of iconoclasts such as Thorstein Veblen, whom Ross sees as an exception proving the rule.

Exceptionalism has proved remarkably resilient in the face of tumultuous changes in American history. Each historical period posed a different set of challenges to the exceptionalist framework, threatening to undermine it. Social scientists beat back these challenges, reformulating their theories to integrate historical developments. Social scientists of the Gilded Era, for example, gradually accepting the inevitability of industrial development, sought to protect American institutions from radical change and the "threat of socialism," by advocating limited reforms and affirming continuous progress within the fixed parameters of the exceptionalist heritage. Progressive Era social scientists also "tried to relocate the exceptionalist ideal in the progressive stream of liberal history" (p. 253). True, conflict existed in America, but it was not structural class conflict but rather a merely temporary conflict attending the transition to an industrial America.

The exceptionalist logic of the social sciences culminated in a scientistic orientation whose elements were atomized individualism, utilitarianism, and positivism. The effect of scientism has been to fetishize the methods of the natural sciences, to rationalize and obscure structural inequality, and to deny the historicity and normative character of scientific theory. Ross is perhaps at her most persuasive in her discussion of the rise of marginalist neoclassical economics. She demonstrates that, in the guise of science, marginalism offered a defense of modern capitalism and blunted the critical potential of the alternative institutional approach. Yet the "instrumental positivism" that came to prevail in sociology is subjected to a similar critique.

Ross is circumspect in relating the cultural patterns of American social science to the institutional context of a decentralized university system, disciplinary autonomy, and growing professionalization. She notes the overwhelmingly Protestant middle-class background of social scientists

and their unconscious, secularized Christian moralism and millennial expectation. She is sensitive to the gender, racial, class, and ethnic biases encoded in their intellectual positions. Many of her portrayals of exemplary figures, though terse, are extremely insightful—her life history of the libertarian economist Frank Knight is just one case in point.

The most problematic aspect of this work is that Ross does not take into account the criticisms leveled against historicism, particularly its relativistic implications and the "hermeneutic circle." Her skill at pointing out the blinkers on American social science is laudable, and she demonstrates a nuanced appreciation of countervailing nonpositivist tendencies. She does not, however, elaborate criteria for distinguishing scientism from science nor is it clear that she wants to. But the absence of such criteria makes for a certain degree of inconsistency because she appears at times to conflate scientism and science and at other times to hold them separate.

The Origins of American Social Science represents a major contribution by a first-rate intellectual historian and constitutes a powerful defence of alternative, nonpositivist, and critical approaches to social science. The analysis is subtle but forceful, measured but firm. Adherents of reigning social-science paradigms are called to take notice, for Ross has rendered a powerful critique of cherished beliefs. This book is destined to be widely read and ranks as one of the best treatments of American social science in its historical context.

Asian Women Workers in Kuwait
NASRA M. SHAH, SULAYMAN S. AL QUDSI and MAKHDOOM A. SHAH

Dual Chain Migration: Post-1965 Filipino Immigration to the United States
JOHN M. LIU, PAUL M. ONG and CAROLYN ROSENSTEIN

The Experience of Refugees in the Australian Labour Market
MARK WOODEN

Sponsorship and Employment Status Among Indochinese Refugees
in the United States
THANH V. TRAN

Refugee Adaptation and Community Structure:
The Indochinese in Quebec City, Canada
LOUIS-JACQUES DORAIS

Africa's Refugee Crisis: State Building in Historical Perspective
CONSTANCE G. ANTHONY

RESEARCH NOTE:
Immigrants, Natives, and the French Welfare State:
Explaining Different Interactions with a Social Welfare Program
JEREMY HEIN

Book Reviews • Review of Reviews
International Newsletter on Migration • Books Received

AJS

CONTENTS

Book Reviews

IN THIS ISSUE

LAUREN B. EDELMAN is assistant professor of sociology and law at the University of Wisconsin—Madison. Her work focuses on the interplay between law and the employment relation and includes studies of organizational response to civil rights law, dispute handling in organizations, and the role of the personnel and legal professions in shaping organizational response to law. She is currently using data from a nationwide study of organizations' employment practices to study how organizations construct the meaning of compliance with civil rights law and how forms of compliance become institutionalized.

ROSS L. MATSUEDA is associate professor at the University of Wisconsin—Madison. His research examines sibling and family models of delinquency, empirical test of rational choice and deterrence theories, models of work and crime, and methods for the analysis of covariance structures. His article in this issue is part of a larger project examining a symbolic interactionist approach to delinquency.

PHYLLIS MOEN is professor of human development and family studies and of sociology at Cornell University. She is currently serving as director of the Life Course Institute at Cornell and is the author of *Working Parents: Transformations in Gender Roles and Public Policies in Sweden* and *Women's Two Roles: A Contemporary Dilemma*.

DONNA DEMPSTER-MCCLAIN is a research associate in the Department of Human Development and Family Studies at Cornell University. She has been the director of the Women's Roles and Well-Being project since 1986 and was also the acting director for the Cornell Nursery School, 1990–91.

ROBIN M. WILLIAMS, JR., is the Henry Scarborough Professor of Social Science, emeritus, at Cornell University and the editor of *Sociological Forum*. He is coeditor with Gerald David Jaynes of the 1989 National Research Council report, *A Common Destiny: Blacks and American Society*. His current interests include comparative ethnic relations and the sociology of war and other collective conflicts.

DAVID LEE STEVENSON is deputy executive director of the National Council on Education Standards and Testing. He received his Ph.D. from Yale University. His major research interests are the comparative study of educational systems, social organization of schools, and social policy. He is currently conducting a national longitudinal study of U.S. secondary schools.

DAVID P. BAKER is associate professor of sociology and research associate at the Life Cycle Institute at the Catholic University of America, Washington, D.C. He has recently completed a comparative analysis of school effects in Japan and the United States. In addition to comparative analysis of school organization and achievement, his research interests include the history of immigration and the institutional growth of the Roman Catholic Church in America and studies of the effect of immigration on European labor markets.

STEVE RYTINA is associate professor of sociology at McGill University. His larger research interests in inequality and injustice have led to a special focus on how methodological choices have influenced substantive interpretation of mobility data.

ROBERT M. HAUSER is the Vilas Research Professor and the Samuel A. Stouffer Professor of Sociology and the director of the Institute for Research on Poverty at the University of Wisconsin—Madison. Along with John A. Logan and others, he is carrying out a 1992 follow-up of the 10,000 members of the Wisconsin Longitudinal Study of social and psychological factors in achievement. He is also currently studying trends and differentials in social mobility, school enrollment, and grade progression in the United States from the late 1960s to the present.

JOHN ALLEN LOGAN is assistant professor of sociology at the University of Wisconsin—Madison. Apart from his collaboration with Robert Hauser and others on the new wave of the Wisconsin Longitudinal Study, he is working on a method for studying occupational outcomes that takes explicit account of resource and preference distributions among individuals and firms.

DAVID B. GRUSKY is assistant professor of sociology at Stanford University. He is also a National Science Foundation Presidential Young Investigator and coeditor of the Westview Press Social Inequality Series. His current research addresses the structure of cross-national similarities and differences in mobility regimes, the evolution of attitudes toward sex-based inequalities, and recent trends in patterns of occupational and geographic mobility. He is editor of the forthcoming book, *Social Stratification: Class, Race, and Gender in Sociological Perspective* (Boulder, Colo.: Westview Press, 1992).

STEPHEN E. VAN ROMPAEY is a doctoral student in the Department of Sociology at Stanford University. His research interests include social stratification, geographic mobility within the United States, and the social demography of Latin American nations.

Information for Contributors

Editorial Procedures: All papers deemed appropriate for the *AJS* are sent out anonymously to readers. The editors rely heavily on the judgments of those readers, although they are not bound by them. To protect anonymity, only the title should appear on the manuscript Attach a cover page with the title of the manuscript, your name, and your affiliation. All identifying references and footnotes should appear on a separate page. A legible, carefully prepared manuscript will facilitate the work of readers. Manuscripts are accepted for publication subject to nonsubstantive editing with the understanding that the *AJS* has the right of first publication. Submission of a manuscript to the *AJS* is taken to indicate the author's commitment to publish in this *Journal*. No paper known to be under consideration by any other journal will be reviewed by the *AJS*. Upon publication, all rights, including subsidiary rights, are owned by the University of Chicago Press. The author retains the right to use his article without charge in any book of which he is the author or editor after it has appeared in the *Journal*

Manuscript Acceptance Policy: While it is our policy to require the assignment of copyright on most journal articles and review essays, we do not usually request assignment of copyright for other contributions. Although the copyright to such a contribution may remain with the author, it is understood that, in return for publication, the journal has the nonexclusive right to publish the contribution and the continuing right, without limit, to include the contribution as part of any reprinting of the issue and/or volume of the journal in which the contribution first appeared by any means and in any format, including computer-assisted storage and readout, in which the issue and/or volume may be reproduced by the publisher or by its licensed agencies.

Preparation of Copy

1. Type *all* copy—including indented matter, footnotes, tables, and references—*double-spaced,* allowing generous margins at top, bottom, and sides of page. Please do not break words at ends of lines or justify right-hand margins. Avoid proportional spacing fonts. Indicate italics by underlining only. Number all footnotes, including acknowledgment footnote (which should be "1"). *Provide a word count on the cover page of the manuscript.*

2. Type each table on a separate page Refer to each table in numerical order in the text. Prepare tables without vertical lines. Tables should not contain more than 20 two-digit columns or the equivalent.

3. Clarify all mathematical symbols (e.g., Greek letters) with words in the margins of the manuscript.

4. Draw figures on white paper with india ink. Original or glossy print of professionally drawn figure will be required if manuscript is accepted.

5. Include a brief abstract (not more than 100 words) summarizing the findings.

6. *Five copies* of the manuscript and abstract, on white paper, must be submitted. Manuscripts should be addressed to the Editor of the AMERICAN JOURNAL OF SOCIOLOGY, 5835 South Kimbark, The University of Chicago, Chicago, Illinois 60637. *Manuscripts are not returned to authors.* Enclose a self-addressed stamped postcard, containing the manuscript title, to obtain acknowledgment of manuscript receipt. No other acknowledgment will be sent.

7. Each article submitted to the *AJS* must be accompanied by a check or money order for $15.00, payable to The University of Chicago Press in U.S. currency or its equivalent by money order or bank draft. Papers will not be reviewed until this fee has been paid. The submission fee will be waived for papers solely by student authors. Submissions from students must be accompanied by a faculty member's certification of their student status. Citizens of countries with restrictions on the export of U.S. dollars may request waiver of the submission fee.

Format of References in Text: All references to monographs, articles, and statistical sources are to be identified at the appropriate point in the text by last name of author, year of publication, and pagination where appropriate, all within parentheses. Footnotes are to be used only for substantive observations. Specify subsequent citations of the same source in the same way as the first one; do not use "ibid.," "op. cit.," or "loc. cit."

1. When author's name is in the text: Duncan (1959). When author's name is not in text: (Gouldner 1963).

2. Pagination follows year of publication: (Lipset 1964, pp. 61–65).

3. For more than three authors, use "et al." For institutional authorship, supply minimum identification from the beginning of the complete citation: (U.S. Bureau of Census 1963, p. 117).

4. With more than one reference to an author in the same year, distinguish them by use of letters (*a, b*) attached to the year of publication: (1965*a*).

5. Enclose a series of references with a single pair of parentheses, separated by semicolons.

Format of References: List all items, including machine-readable data files (MRDF), alphabetically by author (providing the full list of multiple authors) and, within author(s), by year of publication, in an appendix titled "References." Examples follow:

Davis, James Allan. 1978. *General Social Survey, 1972–1978: Cumulative Data* (MRDF). NORC ed Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

Davis, K. 1963*a*. "Social Demography." Pp. 124–37 in *The Behavioral Sciences Today*, edited by Bernard Berelson. New York: Basic.

———. 1963*b*. "The Theory of Change and Response in Modern Demographic History." *Population Index* 29:345–66.

Goode, W. J. 1967. "The Protection of the Inept." *American Sociological Review* 32:5–19.

Moore, Wilbert E., and Arnold S. Feldman. 1960. *Labor Commitment and Social Change in Developing Areas*. New York: Social Science Research Council.

Sanford, Nevitt, ed. 1962. *The American College*. New York. Wiley.

U.S. Bureau of the Census. 1970. *Census of Population and Housing 1970, Summary Statistic File 4H: U.S.* (MRDF). DUALabs ed. Washington, D.C.: Bureau of the Census. Distributed by DUALabs, Rosslyn, Va.

Weber, Max. (1921) 1968. "Society's Problems." Pp. 12–16 in *Economy and Society,* edited by Guenther Roth and Claus Wittich. New York: Bedminister.

(Rev. 7/91)

Legal Ambiguity and Symbolic Structures: Organizational Mediation of Civil Rights Law[1]

Lauren B. Edelman
University of Wisconsin—Madison

Laws that regulate the employment relation tend to set forth broad and often ambiguous principles that give organizations wide latitude to construct the meaning of compliance in a way that responds to both environmental demands and managerial interests. Organizations respond initially by elaborating their formal structures to create visible symbols of compliance. As organizations construct and institutionalize forms of compliance with laws, they mediate the impact of those laws on society. The author uses data from a nationwide survey of 346 organizations to develop models of the creation and institutionalization of organizationally constructed symbols of compliance following the 1964 Civil Rights Act.

INTRODUCTION

As laws regulating organizations grow both in number and complexity, the legal environment becomes an increasingly salient determinant of organizational structure and behavior. Studies of law and organizations measure the impact of law (e.g., Leonard 1986; Jones 1985a, 1985b; Burstein and Monaghan 1986; Heckman and Payner 1989; Beller 1982), criticize the form and structure of law (e.g., Freeman 1982; Bumiller 1988), and identify factors that motivate compliance and noncompliance

(e.g., Hawkins 1984). But our understanding of the *process* by which organizations respond to law is very limited. This omission is critical: through the process of response to law, organizations construct the meaning of compliance and thus *mediate* the impact of law on society.

The opportunity for organizations to mediate law is variable. Laws that contain vague or controversial language, laws that regulate organizational procedures more than the substantive results of those procedures, and laws that provide weak enforcement mechanisms leave more room for organizational mediation than laws that are more specific, substantive, and backed by strong enforcement. Laws that regulate the employment relation tend to set forth broad and ambiguous principles that give organizations wide latitude to construct the meaning of compliance. Such laws set in motion a process of definition during which organizations test and collectively construct the form and boundaries of compliance in a way that meets legal demands yet preserves managerial interests.

In this article I examine the initial stages of organizational response to and mediation of law; that is, I explore the definition and institutionalization of what constitutes compliance. I examine organizational mediation of law in the context of equal employment opportunity and affirmative action (EEO/AA) law as part of a larger project designed to explore changes in organizations' EEO/AA practices and the interplay between organizations and their legal environments. Data for the project comes from a nationwide survey of EEO/AA practices in 346 organizations, which I conducted in 1989. In future research, I will use these data to address the long-term effects of organizations' initial constructions of compliance.

Equal employment opportunity/affirmative action law is perhaps the most important legal encroachment on employers' prerogatives since the labor legislation of the 1930s. It seeks to limit employers' ability to perpetuate social advantage or disadvantage through employment opportunities; in so doing it threatens to constrain traditional managerial prerogatives to choose freely whom to hire, fire, and promote. EEO/AA law consists of a body of antidiscrimination mandates comprising statutes, constitutional mandates, and presidential executive orders. Much of my discussion focuses on the two EEO/AA mandates that have most dramatically changed organizations' legal environments: Title VII of the 1964 Civil Rights Act and Executive Order 11246 (EO 11246), issued by President Johnson in 1965. Title VII precludes employers and labor unions with 15 or more employees or members from discriminating on the basis of race, color, religion, sex, or national origin.[2] Executive Order 11246

---

[2] Originally, Title VII did not apply to public employers, but it was amended by the Equal Employment Opportunity Act of 1972 to include public employers.

requires federal employers and private contractors, subcontractors, and unions doing work under or related to a federal contract of $10,000 or more to refrain from discrimination on the same bases and to engage in affirmative action. The organizational changes I discuss are responses to these specific mandates as well as to the general legal environment, which has become progressively more complex and demanding.[3]

Extant literature on EEO/AA law presents a mixed picture of its impact. Studies that focus on the regulatory process tell us that organizations have a strong capacity to resist legal control and that both the structure of regulation and the culture of the business world may buttress that resistance (Wirt 1970; Stone 1975; Vaughn 1982; Coleman 1984). Others point to corporate influence on the regulatory agencies ("capture") as a source of weak enforcement (Jaffe 1954; Blumrosen 1965; Bullock and Lamb 1984) or emphasize the ineffective use of legal sanctions by government agencies (Wirt 1970; Leonard 1985).

Critical analyses of social reform laws generally, and of EEO/AA law in particular, emphasize that those responsible for formulating, interpreting, and enforcing the law are part of the dominant class and that they use their authority to construct law in a way that preserves the status quo while giving the appearance of change (e.g., Freeman 1982; Tushnet 1984; Bumiller 1988; Crenshaw 1988). In this vein, Bumiller (1988) points out that the reform potential of EEO/AA law is limited because victims of employment discrimination tend not to pursue legal redress. And Freeman (1982, p. 110) charges that "the goal of civil rights law is to offer a credible measure of tangible progress without in any way disturbing the basic class structure."

Empirical studies of the appellate EEO/AA litigation show that employment discrimination plaintiffs are less likely than other types of plaintiffs to win their cases. In a study of cases involving job discrimination brought in federal district courts from 1978 to 1985, Eisenberg (1990) reports that plaintiffs won 21% of their cases. In comparison to most other types of cases, this figure is quite low. Of 72 categories of claims, only four had lower plaintiff win rates than employment discrimination. For example, Eisenberg's data show that the average plaintiff win rate is 47% for personal injury tort cases and 63% for contract cases. Litigation success rates are, however, only one dimension of success; discrimi-

---

[3] During the 1960s and early 1970s, a number of EEO/AA laws were passed and executive orders issued to address the civil rights of various groups. These laws include Executive Order 10925 in 1961, the Equal Pay Act of 1963, the Age Discrimination in Employment Act of 1967, the Rehabilitation Act of 1973, the Vietnam Era Veterans' Readjustment Assistance Act of 1972 and Vietnam Era Veterans' Readjustment Assistance Act of 1974.

nation complaints may be resolved at earlier stages in administrative or legal process.

While accounts that focus on law and the regulatory process find them seriously flawed, studies of work-force demographics show improvements in the economic and occupational status of minorities and women, albeit not necessarily in response to the law. Leonard (1984a, p. 148), for example, reports that between 1966 and 1978, the proportion of nonwhite males in the manufacturing work force grew from 8% to 11%, while the proportion of females increased from 26% to 31%. Minority and female employment shares especially grew between 1974 and 1980 among government contractors (Leonard 1986). Beller (1982) reports a decrease in the sex differentials in earnings and in the sex differentials of being employed in male occupations after 1972. Bound and Freeman (1989) report that the proportion of black men working as professionals or managers relative to the proportion of white males in those categories rose from 32% in 1964 to 64% in 1986. Eberts and Stone (1985) report increases in the promotion rates of women since 1964. Heckman and Verkerke (1990) report that while the median income of nonwhite males was 57% of the median white male income in 1964, it had grown to 66% of the median white male income by 1985. Not all reports are so positive: the black unemployment rate has remained at approximately twice that of whites (Butler and Heckman 1977), and recent studies find that the progress of minorities has slowed and, by some measures, retrogressed since 1980 (Bound and Freeman 1989; Heckman and Verkerke 1990). Still, the overall work-force position of minorities and women is clearly better today than it was in 1964.

The inconsistencies between these accounts can be partially explained by focusing on the process by which organizations respond to law rather than by viewing law and the legal process as a given authority that organizations either obey or resist. Several recent studies in other areas of law point to the importance of viewing law as indeterminate and compliance as emergent. Clune's (1983) analysis of the implementation of administrative law treats compliance as a dynamic and emergent relationship between regulators and the regulated. Hawkins's (1984) analysis of the enforcement of antipollution law shows how the uncertainty of regulatory law engenders the construction of compliance through regulatory processes. And Burk's (1988) study of organizational response to federal securities law show that law may have unintended effects by motivating social actors to use the law as a resource to pursue their own interests; this in turn affects the character of organizations and their environments.

This study views compliance with EEO/AA law as a social process that evolves over time as organizations seek to adapt the law to fit their

own interests. I take seriously the commentary of critical scholars, but I suggest that organizations, rather than resist law overtly, are motivated by the weaknesses of EEO/AA law and the mechanics of the legal process to construct law in a manner that is minimally disruptive to the status quo. My analysis draws upon institutional theories of organizations, which emphasize organizations' dependence upon and interaction with their normative environments (Meyer and Rowan 1977; Meyer and Scott 1983; DiMaggio and Powell 1983; Edelman 1990). In particular it draws upon the legal environment theory that I have developed as a means of explaining organizational response to general legal norms (Edelman 1990), and it extends that theory to address organizational response to direct legal mandates.

In earlier work (Edelman 1990), I have argued that law creates a "legal environment" that consists not only of law and the sanctions that are built into law, but also of societal norms and culture associated with the law. Organizations that are sensitive to their legal environments develop forms of governance that conform to legal norms in order to achieve legitimacy. Over time, some organizational responses to the legal environment diffuse among organizations and become institutionalized. Thus by influencing organizations' environments, law has an important indirect effect on organizational behavior that goes significantly beyond the direct effect of law and legal sanctions.

EEO/AA LAW AND THE LEGAL ENVIRONMENT

Equal employment opportunity/affirmative action law influences the legal environment by changing public expectations about employees' civil rights and providing a basis for criticizing well-ingrained patterns of governance that favor whites and males. The law bolsters efforts by pressure groups to realize civil rights, and it legitimates employees' claims to those rights. By empowering those who make demands upon organizations, EEO/AA law creates a legal environment that encourages organizations to comply, or appear to comply, with EEO/AA law. But normative pressure from the legal environment does not easily erode long-held managerial prerogatives. Employers resist EEO/AA law not only because it infringes upon managerial discretion but also because employers view EEO/AA law as requiring inefficient and irrational business practices.

The conflict between EEO/AA law and managerial interests poses a dilemma to organizations: they must demonstrate compliance in order to maintain legitimacy and at the same time they must minimize law's encroachment on managerial power (Edelman 1990). This dilemma motivates a process of response to law in which organizations test, negotiate, and collectively institutionalize forms of compliance that, to the greatest

extent possible, maximize both interests; it is through this process that organizations "mediate" the law.

THE CONDITIONS OF ORGANIZATIONAL MEDIATION OF LAW

Equal employment opportunity/affirmative action law is especially open to organizational mediation for three reasons: (1) it is ambiguous with respect to the meaning of compliance; (2) as construed by the courts, it constrains organizational procedures more than the outcome of those procedures; and (3) its enforcement mechanisms are relatively weak. Each of these characteristics broadens the scope of organizational behaviors that may be considered compliant.

Ambiguity

Title VII of the 1964 Civil Rights Act, section 703(a), states that

> It shall be an unlawful employment practice for an employer: (1) to fail or refuse to hire or to discharge any individual, or *otherwise to discriminate* against any individual with respect to his compensation, terms, conditions, or privileges of employment, *because of such individual's race, color, religion, sex, or national origin;* or (2) to limit, segregate, or classify his employees or applicants for employment in any way which would deprive or tend to deprive any individual of employment opportunities or otherwise adversely affect his status as an employee, *because of such individual's race, color, religion, sex, or national origin* [Emphasis added]

This language is ambiguous both in a legal sense and with respect to organizational policy. The basic ambiguity concerns the meaning of the phrase "to discriminate" together with the fact that the law uses the general terms "race" and "sex" rather than specifically prohibiting discrimination against "racial minorities" and "women."[4] It is interesting that, although Title VII includes definitions for 11 terms, it does not define the term most central to the law: "discrimination."

The ambiguous language of Title VII leaves the law open to at least two interpretations: (1) a procedural interpretation that emphasizes equality of treatment and thus implies that employment practices and procedures should be carried out in a race- and gender-blind manner and (2) a substantive interpretation that emphasizes equality of outcome and thus would require race- and gender-conscious treatment in order to achieve equity in work-force representation (Belton 1981; Fiss 1974; Free-

---

[4] For the sake of brevity, I use race and sex throughout the paper to refer to all protected categories: race, color, religion, sex, and national origin (which are protected by Title VII) as well as veteran status, age, and other categories that are protected by other laws

man 1982; Blumrosen 1972; Burstein 1990). This ambiguity has generated significant political debate and been the subject of much EEO/AA litigation over the past 25 years. But, more important, the legal ambiguity of Title VII means that employers have little guidance as to what, other than eliminating overtly discriminatory policies and practices, employers must do in order to comply with Title VII.

The only tangible requirements of Title VII are work-force reporting requirements, which do not in and of themselves necessitate any change in employment policies or practices.[5] While the reporting requirement may encourage organizations to hire more minorities or women to make their numbers look good, it does not help to clarify, or indeed even address, what constitutes compliance.

Executive Order 11246 is even more complex than Title VII because it explicitly refers to "affirmative action," an often-used term that has no clear legal meaning (Jones 1985a). Section 202 of EO 11246 states, "The contractor will take *affirmative action* to ensure that applicants are employed and employees are treated during employment *without regard to their race, color, religion, sex, or national origin*" (emphasis added).

Those who embrace the substantive interpretation of Title VII see the "affirmative action" requirement as basically consistent with the requirement that employees be treated "without regard to their race, color, religion, sex, or national origin" (see, e.g., Fiss 1974; Belton 1981), but those who support the procedural interpretation of Title VII contend that the two requirements are contradictory (e.g., Smith 1978).

Federal employers and contractors had little guidance as to the nature of their obligation under EO 11246 until February 1970, when the Office of Federal Contract Compliance (OFCC, later the OFCCP),[6] charged with enforcing EO 11246, issued Order 4 as part of the OFCC Rules and Regulations. The issuance of Order 4 followed a ruling by the Comptroller General that the affirmative action obligation was too vague to meet the legal obligation that minimum contract standards must be clear to the parties. Revised Order 4, which replaced Order 4 in December 1971, requires organizations with 50 or more employees and contracts of $50,000 or more to undertake a "utilization analysis" to determine whether there is an underrepresentation of minorities and women in certain job categories, given their availability in the relevant labor market,

[5] The rules implementing Title VII require organizations with 100 or more employees to report annually to the Equal Employment Opportunity Commission (EEOC), which is charged with enforcing Title VII This report consists of a form that summarizes, by race, gender, and ethnicity, the composition of the work force and the composition of applicants to that work force

[6] The Office of Federal Contract Compliance was renamed the Office of Federal Contract Compliance Programs in 1975

and to establish "goals and timetables" to remedy any deficiencies found through the utilization analysis. But Revised Order 4 incorporates, rather than eliminates, the conflict between the affirmative action and nondiscrimination clauses of EO 11246. Section 60-2.12(e) specifies that *"Goals may not be rigid and inflexible quotas which must be met,* but must be targets reasonably attainable by means of applying every good faith effort to make all aspects of the affirmative action program work" (emphasis added), while sections 60-2.12(g) and (h) state that *"Goals, timetables, and affirmative action commitments must be designed to correct any identifiable deficiencies.* Where deficiencies exist and where numbers or percentages are relevant in developing corrective action, the contractor shall establish and set forth *specific goals* and timetables for minorities and women" (emphasis added).

The meaning of a specific but flexible goal is less than clear. And, like Title VII, although the rules and regulations for compliance with EO 11246 give definitions for 24 other terms, there are no explicit definitions for "discrimination," "affirmative action," or "equal employment opportunity."

Thus the basic ambiguity in EEO/AA law as to what constitutes discrimination and what organizations must (and can) do to comply leaves much open for the courts to interpret. And until judicial clarifications are made, it leaves much for organizations to interpret.

Procedural Emphasis

Laws that constrain procedure more than substance widen the latitude for organizational response because they make it more difficult to detect discrimination that is not blatant. Procedural constraints enhance the potential for organizations to develop forms of compliance that appear to comply with the law but have little substantive effect.

While the Supreme Court has never issued a definitive decision with respect to the conflict between the procedural and substantive interpretations to Title VII, it has—especially since the mid-1970s—been moving closer to the procedural interpretation. In 1971, the Supreme Court appeared to endorse a substantive interpretation of the law in *Griggs v. Duke Power Co.,* when it announced that a finding of discrimination could be based upon the consequences of employer actions rather than the intent underlying them. But since 1976, courts have begun to narrow the applicability of *Griggs* so that Title VII plaintiffs usually must prove intent to discriminate, to broaden the affirmative defenses available to employers, and, more generally, to retreat from "preferential" treatment

for minorities designed to achieve equality in outcome (Belton 1981; Freeman 1982).

Executive Order 11246 seems more result oriented than Title VII since it requires affirmative action (albeit without defining it) and because the rules implementing the order require that contractors establish employment goals for minority and female employment and timetables for achieving those goals. But OFCCP practices and the rules and regulations implementing EO 11246 transform the order into a more procedural constraint by stating explicitly that contractors need not achieve their goals but need only demonstrate a *good faith effort* to do so. The rules state that "no contractor's compliance status shall be judged alone by whether or not it reaches its goals and meets its timetables. Rather, each contractor's compliance posture shall be reviewed and determined by reviewing the contents of its program, the extent of its adherence to this program, and its good faith efforts to make its program work toward the realization of the program's goals within the timetables set for completion" (41 C.F.R., sec. 60-2.15).

The good faith standard helps to reduce the inconsistency between the nondiscrimination and affirmative action clauses of EO 11246 since goals that are imperfectly achieved constitute less of a reverse discrimination threat to whites and males. But it also emphasizes organizational procedures and undermines the likelihood of substantive results.[7] Both Title VII and EO 11246, then, leave open the possibility for organizations to create the appearance of compliance without much change to the racial, ethnic, and gender composition of their work forces.

Weak Enforcement Mechanisms

Weak enforcement mechanisms further obscure the boundaries of compliance by providing inadequate and inconsistent feedback on what organizational practices are legal. The major source of weakness in Title VII enforcement is that, although the Equal Employment Opportunity Commission (EEOC) and the Justice Department initiate a small number of lawsuits, the mobilization of EEO/AA law depends primarily upon complaints initiated by individual victims of discrimination. To obtain remedies for discrimination under Title VII, an individual files a complaint with the EEOC or, in some states, the state fair employment

---

[7] The emphasis on good faith is not unique to EEO/AA law. Hawkins (1984) points out that field officers often define compliance as good faith efforts when enforcing antipollution law. In the case of EEO/AA law, the good faith standard was probably necessary to avoid having the guidelines struck down by the courts as illegal quotas (James E. Jones, personal communication with the author, 1990).

agency. The EEOC may try to conciliate, and, if that fails, may (since 1972) bring a lawsuit in federal court. If the agency does not pursue the complaint, the aggrieved employee's only option for redress is to file a lawsuit individually. In either case, the process is lengthy, expensive, and often oppressive, with no guarantees of redress.

Lawsuits are a primary means of defining the boundaries of compliance, yet research shows that people who see themselves as victimized by employment discrimination rarely pursue legal redress (Bumiller 1987, 1988; Miller and Sarat 1981). Bumiller and others report that victims of employment discrimination often choose not to pursue legal redress because of the cost of litigation, fear of retaliation by employers, the desire to avoid perceptions of victimization, and the victim's own perception of the futility of the legal process (Felstiner, Abel, and Sarat 1981; Miller and Sarat 1981; Carlin and Howard 1965; Bumiller 1988; Galanter 1974; Blumrosen 1965). And, in a survey of individuals' legal problems and the actions they took, Miller and Sarat (1981) found that, whereas on average 5% of all perceived injuries result in court filings, that figure is less than 1% for employment discrimination grievances.

Recent Supreme Court decisions that make it harder for plaintiffs to prevail on Title VII issues, moreover, may further diminish the likelihood that aggrieved employees will pursue legal channels.[8] Employers may play upon employees' reticence to initiate the legal process by creating internal alternatives to legal channels that are easier and less costly for employees to use. While internal channels for challenging discrimination may provide some redress, the handling of employment discrimination claims internally results in fewer judicial opinions on what constitutes compliance.

Compared to Title VII, the enforcement of EO 11246 is less dependent upon individual mobilization: the OFCCP has authority to enforce the order directly and has a number of sanctions available to it. The order authorizes the OFCCP to delay contract awards, to issue a notice requiring a contactor who appears to have violated the order to show cause why enforcement proceedings should not be instituted, to withhold progress payments, and to bar noncompliant organizations from future contracts.

But contract compliance enforcement has been notoriously weak. The

[8] In *Price Waterhouse v. Hopkins,* the Supreme Court held that, where an employee has made a prima facie case of discrimination, the employer may avoid liability by showing that he or she would have taken the same action even in the absence of discrimination. In *Wards Cove Packing, Inc. v. Atonio,* the Court reduced employers' burden of justifying policies that have an adverse impact on protected employees. In *Lorance v. AT&T Technologies, Inc ,* the Court severely limited the time employees have to challenge revisions of seniority systems as discriminatory.

OFCCP (and previously the OFCC) has been repeatedly criticized by the U.S. Commission on Civil Rights (USCCR), the Congress, and studies by independent agencies such as the Brookings Institution for failure to issue guidelines, failure to define "affirmative action," failure to train its staff, and, in particular, failure to use its power to suspend current and bar future government contracts when violations are found (USCCR 1971; House Committee on Education and Labor 1987; Nathan 1969). The Committee on Education and Labor (1987, p. 39) reports that "enforcement powers for the antidiscrimination and affirmative action portions of the Executive Orders were slowly developed, and once instituted, were underutilized."

Enforcement of EO 11246 was virtually nonexistent until the late 1960s (Committee on Education and Labor 1987; Nathan 1969; Jones 1970, 1982, 1985a). Enforcement efforts were somewhat stronger during the 1970s, in part because Order 4 and Revised Order 4 established standards for the evaluation of contractors.[9] Jones (1982, p. 69) characterizes OFCCP enforcement efforts as most enthusiastic from 1972 to 1980 and writes that "the numerical goals and timetables . . . [were] perceived as increasing the likelihood of hearings leading to the imposition of sanctions" (1982, p. 93). But the 1980s saw a significant decrease in OFCCP enforcement activity: the Committee on Education and Labor (1987, p. 16) reported that, since 1980, "effective enforcement has come to a virtual standstill."

The ambiguity and weak enforcement mechanisms of EEO/AA law together with judicial and administrative constructions of law that emphasize fair procedure over fair results weaken the capacity of EEO/AA law to effect change directly. Why, then, do studies of work-force demographics find, fairly consistently, that the status of minorities and women improved significantly between 1964 and 1980, especially among federal contractors? (See Leonard 1984b, 1986.) A partial answer lies in the fact that, because of normative pressure from their legal environments, organizations do not simply ignore or circumvent weak law, but rather construct compliance in a way that, at least in part, fits their interests.

---

[9] There were some enforcement efforts prior to 1970 that influenced the changes in the 1970s. In 1968, the OFCC established "Hometown Plans" for the construction industry in Philadelphia, Cleveland, San Francisco, and St. Louis. The Cleveland and Philadelphia plans were among the earliest compliance programs to use numerical employment goals for the hiring of minorities and women (Committee on Education and Labor 1987). Also in 1968, the OFCC issued regulations for nonconstruction contractors, which, for the first time, required that contractors with 50 or more employees and a contract of $50,000 or more submit a written affirmative action plan (33 Fed. Reg. 7804 (1968); codified in 41 C.F.R. 60-2).

THE PROCESS OF MEDIATION OF LAW

Structural Elaboration

Laws that are ambiguous, procedural in emphasis, and difficult to enforce invite symbolic responses—responses designed to create a visible commitment to law, which may, but do not necessarily, reduce employment discrimination. Organizations respond visibly to law by elaborating their formal structures. I use the term "formal structures" to refer to the configuration of offices and positions and the formal linkages between them (the "organization chart") as well as to formal rules, programs, positions, and procedures. By contrast, informal structures refer to the actual communication channels between offices and positions, the actual behaviors of individuals who occupy them, and informal norms and practices (Scott 1983). The two are not necessarily coupled: informal practices and norms often deviate from formal procedures and rules (Cyert and March 1963).

Organizations' formal structures are more visible to the outside world than their informal structures. As a strategy for achieving legitimacy, organizations adapt their formal structures to conform to institutionalized norms; the structures are symbolic gestures to public opinion, the views of constituents, social norms, or law (Meyer and Rowan 1977). Organizations seek to appear legitimate for a number of reasons: organizations that appear attentive to EEO/AA law are less likely to provoke protest by protected classes of employees within the firm or community members who seek jobs, they are more likely to secure government resources (contracts, grants, etc.), and they are less likely to trigger audits by regulatory agencies. And, if sued, organizations can point to the structural changes as evidence of the nondiscriminatory nature of their policies and practices. While failure to look compliant is unlikely to result in an organization's demise, it does carry an increased risk of legal liability and social disapproval.

Therefore, although EEO/AA law does not specifically require it, organizations respond to law by creating new offices, positions, rules, and procedures (which I will call EEO/AA structures) as visible symbols of their attention to EEO/AA issues and their efforts to comply.[10] To enhance the symbolic value of EEO/AA structures, organizations incorporate the legal language: they create "Affirmative Action offices" or "EEO policies."

[10] For federal contractors, EO 11246 does require that some person be appointed director of affirmative action and that a formal *statement* expressing a commitment to EEO/AA be issued. However, the order does not require a formal rule proscribing discrimination within the firm nor does it require that a special office to handle EEO/AA issues be created.

The ambiguity, procedural emphasis, and weak enforcement mechanisms of EEO/AA law create the conditions under which EEO/AA structures become a source of legitimacy. In the absence of specific substantive requirements (e.g., a requirement that the racial and ethnic composition of the work force reflect that of the surrounding community or available labor pool), visible symbols of attention to EEO/AA law often suffice as evidence of compliance.[11] Hawkins (1984, p. 109) writes that, in the enforcement of antipollution law, "enforcement agents need, as much as concrete accomplishment, some *sign* of compliance. . . . Intention is as important as action. Assessments of conformity tend to be fluid and abstract, rather than concrete and unproblematic" (emphasis in original).

In the arena of EEO/AA law, rules that explicitly prohibit discrimination and offices designed to implement EEO/AA law strongly connote compliance. Given judicial interpretations of EEO/AA law that emphasize fair treatment over fair results, formal policies and procedures that are free of discriminatory language and intent symbolize compliance. Given an enforcement system that depends upon victim mobilization, the effectiveness of these symbols is unlikely to be challenged. And, if challenged, structures that give the appearance of fair treatment make it more difficult for victims of employment discrimination to establish discrimination for legal purposes.

Of course, it is not always the case that structural elaboration is merely symbolic; structural change may be a means of achieving real improvement in minority and female employment status. Nonetheless, structural elaboration does not guarantee change. The symbolic value of EEO/AA structures does not depend strictly on their effectiveness (although affirmative action offices that are ineffective and rules that are repeatedly violated are open to challenge). Thus organizations may, to varying extents, "decouple" EEO/AA structures from other personnel and governance activities (such as hiring and promotion) in order to reduce the extent to which law constrains managerial functions (Weick 1976). Offices may be decoupled by locating them in the bureaucratic hierarchy in a way that reduces their authority. Decoupling should not be confused with autonomy: EEO/AA offices may be given autonomy and authority so that they may monitor organizational practices in other departments,

---

[11] Revised Order 4 explicitly encourages structural elaboration in Subpart C, which is titled "Methods of Implementing the Requirements of Subpart B." Those methods include specific recommendations for compliance, such as using written statements to disseminate EEO/AA policy (publicize it in company newspaper, include nondiscrimination clauses in all union agreements, etc) and appointing a director of EEO programs (who would serve as liaison between contractor and minority organizations, audit training programs and employment practices, and provide career counseling).

but if that monitoring can affect other departments, the office should be considered coupled.

Rules may be decoupled by nonenforcement. Decoupling EEO/AA rules not only helps to preserve traditional managerial prerogatives—it also decreases the need to resolve contradictions and ambiguity inherent in the law. If rules proscribing discrimination are not enforced, the question of whether compensatory measures for minorities constitutes reverse discrimination is less likely to arise.

The power to manipulate formal structure to fit legally prescribed models of proper organizational governance, then, becomes an important means by which organizations resolve the conflict between managerial interests and law. By creating formal structures that are, or appear to be, mechanisms for implementing legal rules, organizations visibly demonstrate their commitment to comply with EEO/AA law. At the same time, the creation of formal structures can minimize law's intrusion on managerial prerogatives since they do not commit organizations to a particular type or degree of compliance. In some organizations, administrators will expect rules to be enforced and affirmative action offices to work actively and enthusiastically to improve the status of women and minorities within the firm. In others, administrators will create EEO/AA structures as substitutes for compliance, as shams.

Structural elaboration is merely the first step in the process of compliance. Once EEO/AA structures are in place, the personnel who work with or in those structures become prominent actors in the compliance process: they give meaning to law as they construct definitions of compliance within their organizations. At that stage, the internal politics of the organization as well as the personal agendas and professional backgrounds of EEO/AA personnel give meaning to legal requirements (Edelman, Petterson, Chambliss, and Erlanger 1991; Burk 1988). But while actors within organizations struggle to construct a definition of compliance, structural elaboration signals attention to law, thus helping to preserve legitimacy.

Although the quest for legitimacy is a primary motivation for structural elaboration, it is not the only one. Structural elaboration is also a method by which organizations rationalize their legal environments, which is especially important in the face of legal ambiguity. Because judicial interpretations of EEO/AA law are complex and continually changing, employers may consider it rational to hire specialized personnel who can monitor legal changes and, when necessary, modify organizational policy. These EEO/AA structures can also be a means of coordinating compliance efforts and of managing documentation and reporting requirements (although many organizations incorporate such functions into their preexistent structures). It may be efficient, especially for a large organiza-

tion, to create an affirmative action office to respond to employees' or community members' claims of discrimination. Federal employers and contractors who must submit complex affirmative action plans may create offices to formulate and administer those plans. The EEO/AA rules may also have rational motivations: top administrators may create EEO/AA rules as a means of controlling discrimination among middle-level managers, thus reducing the likelihood of legal challenges.

Furthermore, structural elaboration facilitates the organizational task of responding to ambiguous law by providing a means of incorporating legal ambiguity without (immediately) construing its meaning. Just as the OFCCP incorporates legal ambiguity into its rules and guidelines, organizations can incorporate legal ambiguity into their EEO/AA structures. Rules can require both nondiscrimination and affirmative action without addressing the possible inconsistency between the two concepts. Rules can prohibit and even sanction discrimination without defining it. Similarly, organizations can create affirmative action offices or staffs without defining affirmative action.

Organizations create EEO/AA structures, then, largely as gestures to their legal environments; these structures are designed to secure legitimacy and minimize the threat of liability. Although organizations may have rational (efficiency-related) motivations for creating EEO/AA offices, the symbolic value of EEO/AA structures motivates organizations to create them even in the absence of any rational motivation for doing so. Indeed, EEO/AA structures, once in place, may pose new technical problems to organizations. Formal rules, for example, may create problems by providing a basis for employees to challenge managerial actions (Selznick 1969). And new positions and offices may enable people to pose serious challenges to managerial interests, since the personnel in such positions are often more committed to EEO/AA law than the administration that hired them (Edelman et al. 1991). Organizations are often willing to incur those costs to control or efficiency in order to gain legitimacy by conforming to legal norms. But, as I note below, conceptions of efficiency are themselves shaped by organizations' normative environments so that, over time, EEO/AA structures that were adopted for legitimacy increasingly appear valuable to efficiency.

The Institutionalization of EEO/AA Structures

Legal change engenders a process of institutionalization whereby new forms of compliance are diffused among organizations and gradually become ritualized elements of organizational governance. Whereas organizations that respond to legal change early devise and test ways to demonstrate compliance, organizations that are slower to respond can copy the

apparently successful compliance strategies of other organizations. As EEO/AA structures become more prevalent in the population, the structures themselves help to constitute the legal environment. Their visibility as symbols of attention to EEO/AA law increases, making it easier for society to question the commitment of organizations that lack such structures. Thus, over time, organizations become increasingly likely to adopt EEO/AA structures.

The personnel and affirmative action professions play a critical role in the institutionalization process (Edelman 1990; Edelman, Abraham, and Erlanger 1992). Through professional journals, conventions, and workshops, these professionals help to convey interpretations of EEO/AA law and models of compliance. In so doing, they not only shape organizations' response to law and the legal environment, but also managerial conceptions of the rationality of the law. By the early 1970s, two major themes had emerged in the professional personnel literature on EEO/AA law. The first is that the demonstration of good faith is critical to compliance and may be a safer strategy of compliance than quota-based hiring or promotion (e.g., Thorp 1973; Marino 1980). The second theme is that the formalization of EEO/AA policy is propitious to management's interests in efficiency and high productivity (e.g., Garris and Black 1974; Froehlich and Hawver 1974). As employers begin to believe that EEO/AA structures have rational as well as symbolic value, their resistance to those structures begins to fade, which contributes to the general institutionalization of EEO/AA structures.

The courts also reinforce the institutionalization of EEO/AA structures when they treat those structures as evidence of good faith. Schultz (1990, p. 1789) finds that, in Title VII cases challenging job segregation by sex, in which employers claim that women lack interest in better jobs, "employers have been able to bolster the lack of interest argument by claiming that they made *special efforts* to recruit women to nontraditional work" (emphasis added). Schultz reports that the types of special efforts that employers claimed, and courts approved,[12] include: affirmative action plans with goals and timetables,[13] a training program for women,[14] and a flyer announcing equal employment opportunity for women and minorities on road maintenance work.[15] Judicial support for EEO/AA structures—together with claims in the personnel literature that tend to

---

[12] Some of the examples of special efforts appeared in a draft of Schultz's article but were omitted from the published version.

[13] For example, EEOC v. Sears, Roebuck & Co. (1986), Movement for Opportunity v. Detroit Diesel Allison Division of General Motors Corp. (1980).

[14] Ste. Marie v Eastern Railroad Association (1981).

[15] Mazus v Dept. of Transportation, Commonwealth of Pennsylvania (1980).

exaggerate judicial support—help to reinforce the diffusion of EEO/AA structures.

Thus once law motivates organizations that are sensitive to their legal environments to devise symbols of compliance, the personnel and affirmative action professions together with the courts help to convey, rationalize, and institutionalize these models of compliance. And as legal definitions of fairness incorporate organizational models of compliance, organizations mediate the law.

HYPOTHESES

General Patterns of Institutionalization

In contrast to a simple compliance/noncompliance model of organizational response to law, the process of organizational mediation that I described above suggests that EEO/AA law should motivate a gradual institutionalization of EEO/AA structures in organizations. This pattern should be evident in the growth rate of such structures: the cumulative risk of creating such structures should increase over time. However, the growth rate is not likely to follow a perfectly regular pattern. Legal change may produce an immediate surge in EEO/AA structures as some organizations react directly to law, followed by a more gradual growth of such structures as their symbolic value takes root. Furthermore, the precise form of the growth rate of EEO/AA structures will be sensitive to changes in political pressure and enforcement efforts.

Since EEO/AA structures vary in their cost (and symbolic value) to organizations, structures that are less costly should spread among organizations more quickly than structures that involve a greater commitment of resources by organizations. I consider two types of EEO/AA structures that differ substantially in the investment of organizational resources that they require: formal rules proscribing discrimination and special offices to handle EEO/AA activities. Whereas EEO/AA rules can be put in place with relatively little cost or effort, specialized EEO/AA offices require space, personnel, and a budget. In general, more costly structures carry more symbolic value, but less costly structures will suffice for most organizations.

Organizations may invest in compliance in other ways also: for example, by creating affirmative action officer positions, affirmative action plans (which are required for government contractors), affirmative action training and recruitment programs, and simply by adding EEO/AA responsibilities to preexisting positions. I focus on EEO/AA offices and rules because they represent different types of voluntary structural elaboration, but I do not claim that organizational attention to these two

1547

structures represents the full extent of organizations' compliance activities.

Variation among Organizations: Innovators and Followers

Organizations vary in their responsiveness and sensitivity to the legal environment (Edelman 1990). Organizations that are most sensitive to the legal environment should create EEO/AA structures earlier and at higher overall rates than organizations that are less sensitive to their legal environments. Furthermore, organizations that are more sensitive to their legal environments should create more complex EEO/AA structures—that is, structures that require greater organizational resources and more visibly demonstrate attention to EEO/AA law. Below I discuss factors that render organizations more sensitive to their legal environments. Each of the factors should render organizations more likely to be innovators (i.e., to create EEO/AA structures earlier and at higher overall rates than other organizations) and more likely to create more complex EEO/AA structures (such as EEO/AA offices).

Proximity to the public sphere.—In this discussion, I use the term "public sphere" to mean the culture surrounding the *federal* state and the federal legal order. There are two components of the dimension of proximity to the public sphere, and I measure each separately. The first is the *sector* of which the organization is a part. The public sector—which consists of government agencies—is clearly the closest to the public sphere. Within that sector, distinctions may be made between federal, state, and local government agencies, with the latter two progressively further from the public sphere. State agencies generally have more contact and interaction with the federal government than do local agencies. However, the difference between state and local government agencies is not likely to be great since both may have their own EEO/AA laws and may establish administrative agencies with authority to process discrimination charges.[16] The private sector—which consists of private corporations—is furthest from the public sphere. There are also intermediate sectors that consist of both public and private organizations: educational organizations and hospitals would fit into this category.

The second component of organizations' proximity to the public sphere, is *administrative linkages to the federal government,* which may be either *hierarchical* (as in the case of federal employers) or *contractual* (as in the case of federal contractors). Either type of administrative linkage subjects organizations to presidential executive orders and gives the

[16] The federal EEOC cooperates with, and often defers to, state agencies in processing complaints. Similarly, state agencies may defer to local agencies.

federal government the formal authority to withdraw money in the interest of public policy. Administrative linkages cut across societal sectors and help to merge the public and private spheres.

Proximity to the public sphere renders organizations more sensitive to the legal environments for several reasons. First, they are subject to more direct regulation in the form of executive orders. Second, they operate in an environment in which rule-based governance, bureaucracy, and notions of citizens' rights are highly institutionalized (Edelman 1990). Third, organizations closer to the public sphere are more visible and subject to a higher degree of public scrutiny than private-sector organizations. The nature of public scrutiny also varies: public-sector and educational organizations are more likely to be evaluated on the basis of their conformity with institutionalized norms, both mandated (e.g., EEO/AA law) and nonmandated (e.g., responsiveness to the public interest). By contrast, the public evaluates private firms by the quality of their products (Scott and Meyer 1983).

Size.—Large organizations are more visible to the public and therefore more likely to respond to societal pressure independently of their proximity to the public sphere. Because of their visibility, moreover, large organizations are more likely to be targeted by enforcement agencies; the increased threat of legal sanctions is likely to motivate structural elaboration. Large organizations are also more likely to create affirmative action offices because the greater number of employees makes the EEO/AA reporting requirements more burdensome. And large organizations enjoy economies of scale; they can more easily allocate resources for EEO/AA structures.

Presence of a personnel (or human resources) department.—If personnel departments serve as windows to the legal environment, as I argued earlier, then organizations with personnel departments should respond to their legal environments faster than organizations without personnel departments. This hypothesis is consistent with research that links the personnel profession to more progressive forms of governance (Jacoby 1985; Baron, Dobbin, and Jennings 1986).

Unionization.—The relationship between unions and EEO/AA goals is complex. In many situations EEO/AA goals conflict with labor-movement institutions such as seniority, arbitration of discrimination-related grievances, union job-referral systems, union-supported job standards, and work rules set by collective bargaining (Whalen and Rubin 1977; Wood 1975). Thus many unions have resisted civil rights laws, and minorities have had to fight to expand the interpretation of unions' duty of fair representation to include employment discrimination (Hill 1975).

On the other hand, a number of labor historians and theorists point to instances in which unions help to protect minorities and women from

employment discrimination (Chamberlain, Cullen, and Lewin 1980; Freeman and Medoff 1984; Newman and Wilson 1981). Union support for EEO/AA goals—especially when unions encourage employees to file complaints and take legal action—increases environmental pressure on organizations to demonstrate compliance with EEO/AA goals, thus making organizations more likely to institute symbolic structures. Unions may be motivated to participate in employers' EEO/AA compliance efforts to maintain control over areas of traditional union concern such as hiring and promotion (Loevi 1973). And, unionized organizations tend to be more formalized and to have constraints on managerial prerogatives already in place; thus, attention to the rights of minorities and women constitutes less of a change than it might in nonunion organizations. Further, unions are themselves subject to Title VII and may, like employers, be eager to demonstrate compliance.

Region.—A number of studies have shown greater effects of federal antidiscrimination policy in the South than in other geographic regions (USCCR 1986; Heckman and Payner 1989; Ashenfelter and Heckman 1976). Heckman and Payner argue that federal law had an especially strong impact upon black employment in traditional manufacturing sectors in the South. They point out that the EEOC targeted Southern textiles and held hearings on the industry in 1966 and 1967 and that those hearings were widely publicized.

Legal experience.—An organization's experience with EEO/AA lawsuits should increase its sensitivity to the normative environment since it places a direct cost on noncompliance. Organizations that have been sued in the past should be especially likely to adopt EEO/AA structures to demonstrate compliance. However, the effect of lawsuits may be weak or nonexistent because it is greatly entwined with an organization's degree of outright resistance to EEO/AA law: those organizations that experience lawsuits have already demonstrated a capacity for ignoring legal pressure.[17]

Summary.—If organizations' response to law is conditioned by their sensitivity to the legal environments, then the factors I have discussed (proximity to the public sphere, size, the existence of a personnel department, and having experienced a lawsuit) should be the primary determinants of the rate at which organizations create EEO/AA structures. However, these hypotheses should predict the rate at which organizations create complex EEO/AA structures, such as EEO/AA offices, more accu-

[17] Presumably, the litigiousness of the legal environment would also affect organizations; an organization in an industry or a geographic region that has experienced highly visible lawsuits would be more likely to create EEO/AA structures as symbols of compliance. Unfortunately, I do not have data to test the latter hypothesis.

rately than the rate at which organizations create structures that can be established more easily, such as EEO/AA rules. Where structures can be created without a substantial expenditure of resources and where they can easily be decoupled from organizational practices, organizations are likely to adopt them even in the absence of high sensitivity to the legal environment.

DATA AND METHOD OF ANALYSIS

The data for this analysis come from a nationwide phone and mail survey of organizational EEO/AA practices, which was conducted during the spring of 1989. The survey was administered to a probability sample of 346 organizations, consisting of 248 private firms, 50 colleges and universities (which I will refer to as "colleges"), and 48 government agencies at the federal, state, and local (county and city) levels. I collected retrospective event-history data on changes in EEO/AA practices, structures, and policies from 1964 to 1989.[18] (See App. A for a more detailed description of the survey and sample.) The variables used in this analysis as well as variable definitions and coding are shown in table 1.

I use event-history analysis to model the rate (also called the hazard rate) at which organizations institute EEO/AA offices and internal EEO/AA rules, respectively; this permits me to examine the process of change over time in response to EEO/AA law (see Tuma and Hannan [1984], Allison [1984], or Carroll [1983] for discussions of event-history analysis). The rate is defined as "the transitional probability over a unit of time where the unit is infinitesimal" (Carroll 1983). Event-history analysis is appropriate for the study of organizational mediation of law because legal environments exert continuous pressure on organizations and organizations are constantly at risk of creating EEO/AA structures. Furthermore, event-history analysis allows for changing values and changing effects of exogenous variables over time, which is important because the normative climate of EEO/AA law in particular and civil rights more generally has changed considerably over the years.

I have event-history data for all of the discrete variables used in this analysis. For size, organizations were asked to give the number of their full-time permanent employees in 1989, 1984, 1975, and 1964. For the dates before 1989, I used regression equations to impute size values where

[18] Because I draw the sample from the population of organizations alive in 1989, there is a potential sample selection problem. However, since there is no clear legal definition of compliance, the lack of EEO/AA structures is very unlikely to engender organizational death; the more likely consequences are failure to secure government contracts and both internal and community disapproval and dissent. Thus, the likelihood of sample selection bias is minimal.

TABLE 1

VARIABLES AND VARIABLE DEFINITIONS

Variables	Definitions
Fixed independent variables:*	
Sector·	
BUSINESS	Private firm
COLLEGE	College or university
FEDERAL	Federal agency
STATE	State agency
LOCAL	Municipal or county agency
SERVICE	Recodes business· "1" if service/"0" if manufacturing
CORE	Recodes business: "1" if core industry/"0" if periphery
SOUTH	"1" if South/"0" if other
PERCSAL	Percentage of full-time permanent employees that are salaried (1989)
Changing independent variables·	
LOG SIZE	Log number of full-time permanent employees (measured in 1964, 1975, 1984, 1989)
FLAG (size imputed)	"1" if size for that time period is imputed/"0" otherwise
FEDERAL CONTRACTOR†	Federal contractor subject to OFCCP regulations
PERSONNEL DEPT†	Department with personnel function
UNION†	Unionized or partially unionized
LAWSUIT†	Has been sued for discrimination-related complaint
Changing dependent variables	
EEO/AA OFFICE†	Separate EEO/AA office
RULES†	Formal rule proscribing discrimination (not just a policy statement)

* Fixed independent variables are independent variables that have one value for an entire period
† This variable was collected as an event history. It is coded "0" before the event occurs and "1" after the event occurs.

those values were missing. Thus, in all analyses using size, I computed a dummy variable (FLAG) to control for bias due to imputation.

To model the formation rates of EEO/AA offices and EEO/AA rules, I use exponential piecewise nonproportional models, which provided significant improvements in fit over proportional models in all analyses. The piecewise specification divides the period from 1964 to 1989 into four time periods to account for variation in the rates of structural elaboration over historical time; both the rates and the effects of explanatory variables are allowed to vary between the specified time periods but are constrained to be constant within time periods. The four time periods,

which were chosen to capture changes in the legal environment and were based on empirical tests, are: 1964, 1965–69, 1970–79, and 1980–89. The initial one year period (January 1–December 31, 1964) captures the start-up period during which the law was enacted but not yet effective.[19] This one-year start-up period was provided to encourage employers to comply voluntarily with the act; thus, this short time period should capture the initial wave of organization response to the new law. During the 1965–69 period, enforcement of EEO/AA law was relatively weak. The decade of the seventies saw the strongest attention to EEO/AA law, while the Reagan era, beginning in 1980, saw a decline in enforcement of EEO/AA law. Empirical tests show that these four time periods also fit apparent historical time-dependence in the rates of office-creation over time.[20]

The models can be represented as:

$$r_p = \alpha_p e^{\beta X + \lambda_p Z_p}$$

where r_p represents the formation rate from the starting state (no structure) to the ending state (the creation of an EEO/AA structure) during period p, α_p represents the constant term for period p, βX represents a

[19] Five organizations created EEO/AA offices prior to the enactment of Title VII, presumably in anticipation of it or in response to state fair employment laws. In a historical analysis such as this, it is often necessary to be somewhat arbitrary in choosing a date to begin "observing" the process. The enactment of Title VII is an important point in the legal history of employment and thus seems to be the most appropriate start date. But since the first civil rights executive order was issued by Franklin D. Roosevelt in 1941 (EO 8802) it is not surprising that a few organizations created EEO/AA structures prior to 1964. The effects of explanatory variables on office creation in the first period, then, should be interpreted as affecting structural elaboration either in anticipation of, or immediately following, the enactment of Title VII.

[20] Although there is no clear best method for choosing time periods, it would be problematic to choose them based *only* on empirical tests because apparent time dependence in the rates may be due to heterogeneity among sample organizations. Therefore, I used empirical tests to ensure that the time periods chosen for theoretical reasons were reasonably consistent with apparent historical time dependence in the data. To determine the best specification of the time periods empirically, I first plotted integrated hazard rates over historical time in order to determine the nature of time dependence in the rates of structural elaboration. I then used χ^2 likelihood-ratio tests to compare several specifications that would be consistent with the integrated hazard plots. Time dependence in the rates of office formation fit the theoretical choice of time periods very well. Time dependence in the rates of rule formation was most consistent with a three-period model: 1964, 1965–76, and 1977–89. However, for theoretical reasons and for consistency, I use the same four period models for both analyses. The only notable difference between the two time-period specifications for rule formation is that in the three-period model, LOG SIZE is statistically significant during the 1965–76 period.

vector of time-independent explanatory variables and their coefficients, and $\lambda_p Z_p$ represents a vector of time-dependent explanatory variables and their coefficients.

RESULTS

Patterns of Structural Elaboration over Time

Overall, organizations are clearly more likely to create symbolic structures that require fewer organizational resources and can more easily be decoupled from actual practices. Table 2 shows descriptive statistics for the creation of EEO/AA offices and rules.[21] In the survey, "rules" were defined as "a written statement or set of rules about procedures to be used in hiring, firing, promoting, or managing employees to ensure that employees are not subject to discrimination." Respondents were told *not* to include mere statements that the organization is an EEO/AA employer. "Offices" were defined as "a separate department for matters relating to equal employment opportunity or affirmative action." By 1989, 297 of the 346 organizations (86%) in the sample had created rules prohibiting discrimination. By contrast, only 64 organizations (18%) had created EEO/AA offices. The majority of organizations (282) incorporated EEO/AA activities into existent offices (usually personnel) rather than creating a new office.

Figure 1 shows integrated hazard plots for the creation of EEO/AA offices over historical (calendar) time; figure 2 gives the same information for the creation of EEO/AA rules. The hazard plots take into account the numbers of organizations at risk of creating EEO/AA structures during each period.[22] Both integrated hazard plots (as well as parts B and D of table 2) are consistent with my argument that there is an initial wave of structural response following a change in the legal environment

---

[21] The dates of rule creation are missing for 58 organizations because they did not respond to the mail portion of the survey. Therefore, for analyses using dates of rule creation, the sample size was 288, and rules were created in 239 cases. The effect of this missing data on the hazard analyses is to produce more conservative estimates of the rate of rule creation, since all of the omitted cases had created EEO/AA rules.

[22] The integrated hazard function provides a nonparametric estimate of the integral of the hazard rate. The area under the curve reflects the cumulative probability over time that organizations will create affirmative action offices (or rules). The slope of the line gives a nonparametric estimate of the rates of office and rule formation over time. The upper and lower lines show the 95% confidence intervals. The integrated hazard is calculated, using the Kaplan-Meier (1958) estimator, from the proportion of organizations at risk of experiencing an event at any point in time that do in fact experience that event; an organization is at risk if it exists at the moment immediately prior to the occurrence of the event.

TABLE 2

DESCRIPTIVE STATISTICS ON THE CREATION OF EEO/AA STRUCTURES

	Government	College	Business	Total
	A. No. (%) of organizations that created EEO/AA offices by 1989			
Office..	19 (39.6)	15 (30.0)	30 (12 1)	64 (18.5)
No office....	29 (60.4)	35 (70.0)	218 (87.9)	282 (81.5)
Total .	48 (100.0)	50 (100.0)	248 (100.0)	346 (100 0)
	B. The creation of EEO/AA offices by time period			
Time period·				
1964	1 (5.3)	1 (6.7)	5 (16 7)	7 (10.9)
1965–69..	2 (10.5)	1 (6 7)	4 (13.3)	7 (10.9)
1970–79..	14 (73.7)	10 (66.7)	15 (50.0)	39 (60.9)
1980–89..	2 (10.5)	3 (20 0)	6 (20.0)	11 (17 2)
Total...	19 (100.0)	15 (100.0)	30 (100.0)	64 (100.0)
	C No (%) of organizations that created EEO/AA rules by 1989			
Rule........	42 (87 5)	40 (80.0)	215 (86.7)	297 (85 8)
No rule ...	6 (12 5)	10 (20.0)	33 (13 3)	49 (14.2)
Total..	48 (100.0)	50 (100 0) . .	248 (100.0)	346 (100.0)
	D. The creation of EEO/AA rules by time period*			
Time period:				
1964.... ..	2 (5.6)	0 (0 0)	14 (8.2)	16 (6.7)
1965–69 .	2 (5.6)	5 (15.6)	23 (13 5)	30 (12.6)
1970–79 .	22 (61.1)	20 (62.5)	76 (44.4)	118 (49 4)
1980–89.	10 (27.8)	7 (21.9)	58 (33.9)	75 (31.4)
Total..	36 (100.0)	32 (100.0)	171 (100.0)	239 (100.0)

* Dates of rule creation were not available for 58 organizations.

and that this wave is followed by a gradual institutionalization of those structural responses over time.[23]

The pattern of expansion appears to be somewhat different for EEO/AA offices than for EEO/AA rules. The rate of *office* formation begins to increase around 1970 and decreases somewhat after 1980. The rise in the rate of EEO/AA office formation, moreover, appears to follow the issuance of OFCC guidelines early in 1970 and corresponds to the period of relatively strong enforcement of EO 11246 during the 1970s. The rate of *rule* formation is rather constant through about 1975 and then remains elevated throughout the 1980s.

[23] Because five organizations that created offices before 1964 are treated as having created offices in 1964, the integrated hazard from 1964–65 is biased upward.
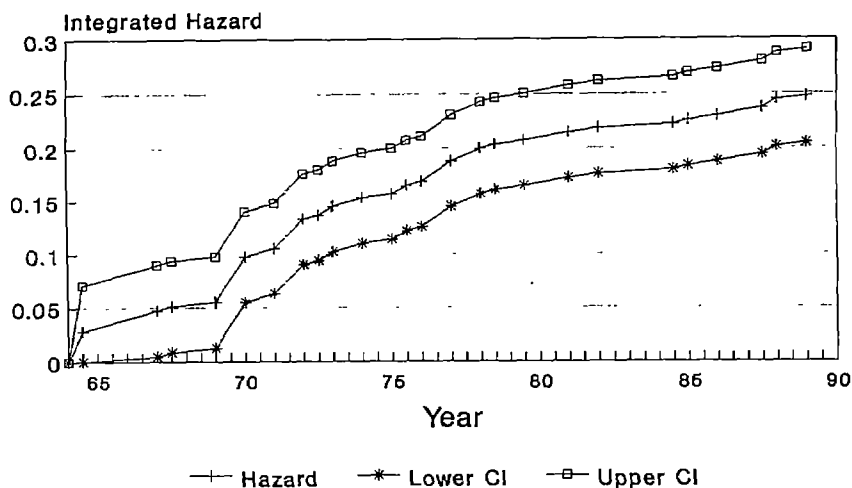
FIG. 1.—EEO/AA office creation, integrated hazard plots, with a 95% confidence interval (CI) (sample size is 346; 64 establishments created offices).
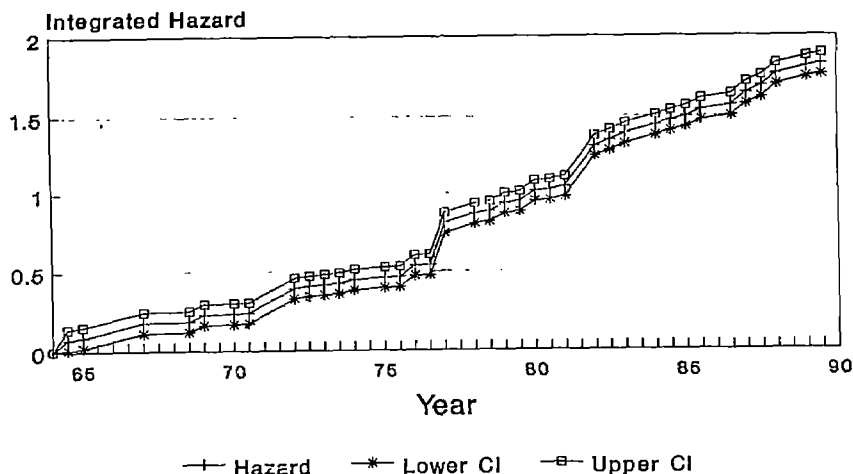
FIG. 2.—EEO/AA rule creation, integrated hazard plots, with a 95% confidence interval (CI) (sample size is 288 with 58 cases omitted [dates of rule creation were not available]; 228 establishments created rules).

Although not conclusive, the fact that the rise in the rate of office creation precedes the rise in the rate of rule creation is consistent with my argument that, as EEO/AA offices become institutionalized symbols of compliance and lend legitimacy to the organizations that have them, organizations that initially resisted legal pressure become more likely to create at least minor symbols of attention to EEO/AA law. Furthermore, the data suggest that organizations that respond early to EEO/AA law may be more likely to create high-investment EEO/AA structures (here, EEO/AA offices), a claim that is further supported by findings I report below.

Sources of Variation in Structural Elaboration

1. *EEO/AA offices.*—EEO/AA offices serve as powerful symbols of attention to EEO/AA law because they signify an allocation of resources, space, and personnel and a commitment to compliance. In fact, the offices are generally quite small. Government agencies have the largest EEO/AA offices, with a mean of 7.1 full-time salaried employees, whereas colleges and business organizations have an average of two or fewer full-time salaried employees. In all but four of the 64 organizations that created EEO/AA offices, there was a personnel office previously in place, and EEO/AA offices are almost always smaller than personnel offices.

The data show that organizations most sensitive to the legal environment are far more likely to create such offices. Table 3 shows seven models of the rate of EEO/AA office formation. All except model 1 (the null model) are piecewise exponential models where the rates are allowed to vary between (but are constant within) the four time periods identified above (1964, 1965–69, 1970–79, and 1980–89).[24] Exogenous variables that have a relatively constant effect on the rate over time are included in a time-independent vector; variables that have a changing effect are allowed to vary between time periods.[25] Although the rate of office creation changes significantly after the initial one-year period, justifying the

[24] Simple exponential models (without time periods) were also run. In these models, only the sector variables and organizational size significantly affected the rates of office formation. However, the piecewise models reveal that the signs of UNION, FEDERAL CONTRACTOR, and PERSONNEL DEP change over time; thus the significant effects of these variables are canceled out in the simple exponential models, which treat the effects of these variables as constant over time. Furthermore, χ^2 likelihood-ratio tests show that the piecewise models provide a substantial improvement in fit over the simple exponential models.

[25] Separate tests were conducted to determine whether the effect of each variable changed significantly over time or could be constrained to be constant across time periods.

TABLE 3

MODELS OF EEO/AA OFFICE FORMATION

Model	Time-independent Variables*	Time-dependent Variables†	Log $\mathcal{L}$	$\chi^2(df)$
1	Null (simple exponential)		−364 3	.
2	Baseline (piecewise)		−346 8	34 86 (3)
3	Sector FEDERAL STATE/LOCAL‡ COLLEGE (BUSINESS)		−329 1	70 26 (6)
4 . ..	Sector FEDERAL STATE/LOCAL COLLEGE (BUSINESS)	LOG SIZE FLAG (size imputed) UNION FEDERAL CONTRACTOR	−310 2	108.15 (18)
5 .	Sector FEDERAL STATE/LOCAL COLLEGE (BUSINESS)	LOG SIZE FLAG (size imputed) UNION FEDERAL CONTRACTOR PERSONNEL DEPT	−307 6	113 35 (21)
6. . ..	Sector FEDERAL STATE/LOCAL COLLEGE (BUSINESS)	LOG SIZE FLAG (size imputed) UNION FEDERAL CONTRACTOR PERSONNEL DEPT LAWSUIT	−306.4	115.77 (24)
7.. .	Sector FEDERAL STATE/LOCAL COLLEGE (BUSINESS) SOUTH SERVICE	LOG SIZE FLAG (size imputed) UNION FEDERAL CONTRACTOR PERSONNEL DEPT	−307 5	113 49 (23)

NOTE.—All models except the null model are four-period piecewise exponential models. The time periods are 1965, 1965–69, 1970–79, 1980–89 Time-dependent variables are entered in each time period but are based on tests of differences across time periods, their values are constrained across periods 1 and 2.

* The coefficients of time-independent variables were constrained to be constant across the four periods. This constraint was based on tests showing that allowing the coefficients of these variables to vary across time periods did not improve the fit of the model.

† The coefficients of time-dependent variables were allowed to vary across time periods Previous tests showed that the effects of these variables were not constant over time

‡ Chi-square likelihood-ratio tests also showed that state and local agencies did not differ significantly from one another Thus their effects are constrained to be equal in all of the models shown.

specification of a new time period in 1965, chi-square likelihood-ratio tests showed that constraining the effects of the time-dependent variables to be constant across periods 1 and 2 improves the fit of the model, albeit only slightly. Chi-square likelihood-ratio tests also showed that state and local agencies did not differ significantly from one another. Thus their effects are constrained to be equal in all of the models shown. Model 4, which includes organizational sector variables (FEDERAL, STATE/ LOCAL, COLLEGE) as fixed over time, and LOG SIZE, UNION, and FEDERAL CONTRACTOR, provides the best fit.[26] However, although model 5 does not provide a statistically significant improvement in fit over model 4, PERSONNEL DEPT is important theoretically and it is statistically significant in the 1970–79 period; thus model 5 may better represent the process of structural elaboration.[27]

Models 6 and 7 in table 3 include several variables that are conspicuous for their lack of effect. Model 6 shows that including LAWSUIT does not provide a significant improvement in fit over model 5. Although LAWSUIT approaches significance in the 1970–79 period, it has no effect in other periods. Since there were relatively few lawsuits before 1970, the lack of an effect in the first period is not surprising. The insignificance of the effect after 1980 may be due to the more conservative judicial stance of EEO/AA law during the Reagan administration. Model 7 shows that including SOUTH does not provide an improvement in fit over model 5. This suggests that the stronger enforcement efforts in the South, which Heckman and Payner (1989) find important, do not increase the rate of office formation. Model 7 also shows that distinguishing between manufacturing and service business firms (SERVICE) does not improve the fit of the model. Two other variables, not shown in table 3, also had no effect on the rate of office formation. These were the measures CORE and PERCSAL.[28]

The coefficients, antilogs, and t-values for the variables in models 4 and 5 are shown in table 4. The coefficients give the effect of the variable on the log of the rate of office formation. The antilogs of the coefficients,

[26] The χ^2 given for each model represents the improvement over the baseline model with no variables and no specification of time periods. Nested models may be compared by using a χ^2 likelihood-ratio test. The test statistic is twice the positive difference between the log-likelihoods of the two models, which under the null hypothesis of no difference will have an asymptotic χ^2 distribution (Allison 1984).

[27] A model that included PERSONNEL DEPT only in the 1970–79 time period would provide a significant improvement over model 4; however, for consistency, I include variables with time-dependent effects in all time periods.

[28] Because PERCSAL was measured only in 1989, including it in the event-history models would require an assumption that the percentage of salaried employees did not change over time. Because I have no evidence to support that assumption, I omit PERCSAL from the models shown in table 3.

TABLE 4

DETERMINANTS OF EEO/AA OFFICE FORMATION

VARIABLES	MODEL 4			MODEL 5		
	Parameter	Antilog	t-value*	Parameter	Antilog	t-value*
Time-independent variables						
Sector						
FEDERAL	2 77	15.92	5.55	2 68	14.68	5.47
STATE/LOCAL†	1.45	4.26	3.93	1.48	4.38	4.01
COLLEGE	1 14	3 14	3.53	1.24	3.45	3.79
BUSINESS	..	..	.	..	..	..
Periods 1 and 2:‡						
Constant, period 1 (1964)	−5.51	.004	−5.01	−5 58	.004	−5.25
Constant, period 2 (1965–69)	−7 84	0004	−7.22	−7.93	0004	−7 52
LOG SIZE	.12	1.13	.77	01	1.11	.68
FLAG (size imputed)	02	1 22	.35	.10	1.10	17
UNION	1.63	5.12	2.75	1.54	4.65	2.55
FEDERAL CONTRACTOR	1 22	3 38	1.90	1 06	2 89	1.60
PERSONNEL DEPT				.62	1.85	.95

Period 3, 1970–79:						
Constant	−6 47	002	−10 27	−6.53	.0014	−7 52
LOG SIZE	.22	1.25	2.20	.16	1.17	1.56
FLAG (size imputed)	.08	1.09	.26	005	1.005	.02
UNION	.46	1.58	1 31	.44	1.55	1.24
FEDERAL CONTRACTOR	.65	1 91	1.81	56	1.75	1.58
PERSONNEL DEPT				.78	2.18	1.94
Period 4, 1980–89:						
Constant	−8.19	1 30	−6 29	−8.26	0003	−6 46
LOG SIZE	.52	1.67	2.38	.56	1 75	2.51
FLAG (size imputed)	−.42	66	−.66	−.39	.68	− 61
UNION	−1.31	.27	−1.62	−1.29	.27	−1 60
FEDERAL CONTRACTOR	−.37	.69	−.56	− 33	.72	−.48
PERSONNEL DEPT				−.34	.71	−.49

* Using a two-tailed test, t-values greater than 1.64 are significant at the .1 level, t-values greater than 1.96 are significant at the .05 level, and t-values greater than 2.57 are significant at the .01 level.

† Chi-square likelihood-ratio tests also showed that state and local agencies did not differ significantly from one another. Thus their effects are constrained to be equal in all of the models shown.

‡ The rate is allowed to vary between the two periods, but the variables coefficients are constrained to be constant.

which give the multipliers of the base rate, are especially informative because they show the relative rates among types of organizations.

The most dramatic differences are among organizations in different sectors, and those differences are relatively constant over time. Based on model 5, which takes into account the effect of personnel departments, colleges created EEO/AA offices at over 3 times the rate of businesses, state and local agencies at over 4 times the rate of businesses, and federal agencies at over 14 times the rate of businesses.[29] The relative numbers of EEO/AA offices created by organizations in the different sectors during each time period can be seen in table 2 above. Overall, 19 of the 48 government agencies (39.6%), 15 of the 50 colleges (30.0%), and 30 of the 248 businesses (12.1%) had created EEO/AA offices by 1989. Furthermore, in the early period after the enactment of Title VII, organizations with contractual or administrative linkages to the federal government created EEO/AA offices at almost three times the rate of other organizations, which supports the argument that linkages to the public sphere, as well as sectoral proximity to the public sphere, affect organizations' response to the legal environment.[30]

Before 1970, unionization is the only factor other than proximity to the public sphere that has a statistically significant effect on the creation rate of EEO/AA offices. From 1964 to 1970, organizations with unions created EEO/AA offices at over four times the rate of organizations without unions. Thus unions appear to expedite organizations' response to EEO/AA law. As stated earlier, this effect may mean that unions are trying to preserve control over traditional areas of union concern, that unions reinforce employers' efforts to create structures that demonstrate compliance because they are also subject to Title VII, or that organizations with unions are likely already to have many formal rules and structures in place, which facilitates the creation of new structures.

The effects of having a union and of being a federal contractor decline over time. This trend reflects a diffusion of EEO/AA offices: by 1970, organizations without federal contracts and unions were becoming subject to greater normative pressure to recognize EEO/AA rights.[31] This

[29] The rates for state agencies are not significantly different from those of local (county and municipal) agencies and thus are constrained to be the same The constraint improves the fit of the model.

[30] I include the eight federal agencies in the sample with federal contractors since both have linkages to the public sphere (and both are subject to EO 11246). However, the effect of FEDERAL CONTRACTOR remains significant when the eight federal agencies are omitted from the analysis.

[31] Although I posit a diffusion of EEO/AA structures among organizations as part of the institutionalization process (since an increasing number of such structures in the population increases environmental pressure to demonstrate compliance), I do not

decline is also in part due to a selection effect: within the sample, organizations with those characteristics that were going to create an EEO/AA office had mostly done so by 1970. Given that federal enforcement, especially of EO 11246, was quite weak until 1970, the data suggest that it is the indirect normative effect of law—rather than the direct threat of legal sanctions—that motivates organizations most sensitive to their environments to create symbols of attention to law.

The effect of organizational size shows the opposite trend. Controlling for the effect of personnel departments, size approaches statistical significance during the 1970–79 period and is significant during the 1980–89 period, while the magnitude of the size effect increases over time. Size alone, then, appears less likely than proximity to the public sphere or unionization to render organizations sensitive to their legal environments; large organizations respond only after the symbolic structures have become more institutionalized. The strength of the size effect during the lax enforcement of the 1980s (and the lack of a size effect earlier in the process), moreover, tends to discredit the argument that offices are created primarily to coordinate the compliance effort. Rather, it suggests that, as symbolic structures become more institutionalized, larger organizations—which are more visible—create offices as symbols of compliance. Further, the statistical significance of PERSONNEL DEPT in the 1970s suggests that personnel officials play some role in promoting the diffusion of symbolic structures, especially when there is more government pressure to comply.

These findings—in particular the strong effects of SECTOR and FEDERAL CONTRACTOR—provide support for my argument that, with respect to high-investment EEO/AA structures, differential sensitivity to the legal environment explains variation in organizations' structural response to law. As shown in the next section, however, this pattern of response is much weaker for structures whose creation involves relatively little cost.

2. *EEO/AA rules.*—In contrast to EEO/AA offices, EEO/AA rules are easy to create and do not require much expenditure of organizational resources; thus the overall formation rate of EEO/AA rules is significantly higher than that of EEO/AA offices. Table 5 shows seven model

---

explicitly model that diffusion. Such a model is not necessary since, given the sectoral and geographic diversity of the organizations in my sample, it is extremely unlikely that there is any direct copying of EEO/AA structures from one sample organization to another. Furthermore, the process I describe is one of organizational response to the general legal environment, which includes laws, public policy, societal norms, and the culture surrounding the law; the practices of other organizations are only part of that legal environment.

TABLE 5

MODELS OF EEO/AA RULE FORMATION

Model	Time-independent Variables*	Time-dependent Variables†	Log $\mathscr{L}$	χ^2 (df)
1 ..	Null (simple exponential)		−873 3	.
2	Baseline (piecewise)		−849 6	47 39 (3)
3.. .. .	FEDERAL CONTRACTOR PERSONNEL DEPT	LOG SIZE FLAG (size imputed)	−818 2	110.24 (13)
4.... .	FEDERAL CONTRACTOR PERSONNEL DEPT Sector FEDERAL COLLEGE STATE/LOCAL‡ (BUSINESS)	LOG SIZE FLAG (size imputed)	−816 3	114 08 (16)
5 .	FEDERAL CONTRACTOR PERSONNEL DEPT	LOG SIZE FLAG (size imputed) EEO/AA office LAWSUIT	−815 2	116 21 (21)
6 . .	FEDERAL CONTRACTOR PERSONNEL DEPT	LOG SIZE FLAG (size imputed) EEO/AA office LAWSUIT UNION	−812 9	120 71 (25)
7 . ..	FEDERAL CONTRACTOR PERSONNEL DEPT	LOG SIZE FLAG (size imputed) EEO/AA office LAWSUIT UNION SERVICE SOUTH	−806 6	133 36 (33)

NOTE.—All models except the null model are four-period, piecewise exponential models The time periods are 1964, 1965–69, 1970–79, 1980–89 Time-dependent variables are entered in each time period

* The coefficients of time-independent variables were constrained to be constant across the four periods. Ths constraint was based on tests showing that allowing the coefficients of these variables to vary across time periods did not improve the fit of the model

† The coefficients of time-dependent variables in this column were allowed to vary across time periods Previous tests showed that the effects of these variables were not constant over time.

‡ Chi-square likelihood-ratio tests also showed that state and local agencies did not differ significantly from one another. Thus their effects are constrained to be equal in all of the models shown.

specifications for the creation of rules proscribing discrimination. These are again, with the exception of model 1, exponential piecewise models that use the same four time periods. The most notable difference from the models of office formation is that the organizational sector variables shown in model 4 do not improve the fit of the model. Models 5 and 6 show, respectively, that neither LAWSUIT nor UNION affects the rate of rule formation. Model 7 shows that neither SERVICE nor SOUTH affects the rate of rule formation. The best-fitting model is model 3,

which includes only three explanatory variables: FEDERAL CON-
TRACTOR, PERSONNEL DEPT, as time-independent variables, and
LOG SIZE as time-varying.[32]

The coefficients for model 3—and for the sake of comparison, for
model 4—are shown in table 6. As can be seen in model 4, the sector
effects on the rate of rule creation are small and insignificant, although
state and local government agencies appear somewhat more likely than
organizations in other sectors to create rules. The cross-sectional data in
table 2 above show that the percentages of organizations that created
rules are similar across sectors: 42 of the 48 government agencies (87.5%),
40 of the 50 colleges (80%), and 215 of the 258 businesses (86.7%) created
rules. However, the effect of FEDERAL CONTRACTOR is statistically
significant if not dramatic: organizations with such linkages create EEO/
AA rules at 1.34 times the rate of other organizations, and this effect
does not diminish over time. These results suggest that OFCCP require-
ments and/or enforcement efforts rather than proximity to the public
sphere per se motivate rule formation.

The strongest determinant of rule formation, however, is PERSON-
NEL DEPT: organizations with personnel departments create rules at
almost twice the rate of other organizations. This strong effect supports
my argument that these departments act as "windows" to the legal envi-
ronment and are critical to the institutionalization of structural responses
to law. However, personnel departments appear to play a greater role in
the formation of antidiscrimination rules than in the creation of offices,
which suggests that personnel departments help to institutionalize lower-
investment symbols of compliance.

Although the presence of EEO/AA offices might be expected to have
an effect on rule formation, the data show no such effect. A closer look
at the order in which organizations created the two structures explains
this finding. Of the 64 organizations that created EEO/AA offices, 5
never created EEO/AA rules, 12 created offices and rules at the same
time, 16 created rules first, 18 created offices first, and, in 13 cases, the
date at which the rules were created was missing. Thus, in this sample,
the existence of one structure does not affect the formation rate of the
other.

Beyond the effects of personnel departments and federal contractor
status, there is not a clear pattern to rule formation. Organizational size
(LOG SIZE) is statistically significant only in the 1970–79 period, and

---

[32] As with office formation, PERCSAL had no effect on the rate of rule formation.
Due to the lack of event-history data on this variable, I do not include it in the models
shown in table 5.

TABLE 6

DETERMINANTS OF EEO/AA RULE FORMATION

Variables	MODEL 3			MODEL 4		
	Parameter	Antilog	t-value*	Parameter	Antilog	t-value*
Time-independent variables:						
Constant						
FEDERAL CONTRACTOR	30	1 34	2 12	31	1 37	2.19
PERSONNEL DEPT	65	1.92	4.42	.64	1.90	4 31
Sector						
FEDERAL				15	1 17	.41
STATE/LOCAL†				.40	1.49	1.87
COLLEGE..				− 07	93	− 36
BUSINESS....				.	.	..
Period 1, 1964:						
Constant	−4.04	.002	−2.93	−4.22	.01	−3.06
LOG SIZE	.97	2.64	1.36	1 05	2.87	1 48
FLAG (size imputed).... ...	−.08	.93	−.16	− 07	.93	− 14
Period 2, 1965–1969:						
Constant	− 3 78	02	−6 72	−3.97	02	−6.94
LOG SIZE	.03	1 03	.10	12	1.12	36
FLAG (size imputed)	− 44	.64	−1 19	− 44	.65	−1.17
Period 3, 1970–79:						
Constant	−3.77	.02	−10.27	−3.92	.02	−10.01
LOG SIZE	52	1 69	2 52	.59	1.80	2.77
FLAG (size imputed)	− 30	.74	−1.65	− 30	.74	−1.63
Period 4, 1980–89:						
Constant	−2.80	.06	− 7 20	−2.89	06	−7.12
LOG SIZE	.12	1.13	52	.17	1.18	.70
FLAG (size imputed)	− 09	91	− 38	−.11	.89	−.48

* Using a two-tailed test, t-values greater than 1 64 are significant at the 1 level, t-values greater than 1 96 are significant at the 05 level, and t-values greater than 2 57 are significant at the 01 level

† Chi-square likelihood-ratio tests also showed that state and local agencies did not differ significantly from one another Thus their effects are constrained to be equal in all of the models shown

the magnitude of the effect is relatively small.[33] The relative lack of differentiation between types of organizations in the rates at which they create EEO/AA rules, together with the fact that by 1989, 300 of the 350 organizations had created rules proscribing discrimination, suggests that when structures that symbolize compliance are easy to institute (and easy to decouple from actual practices), the response is almost universal and the structure becomes highly institutionalized. Indeed, the widespread implementation of EEO/AA rules by 1989 makes organizations that do not have them appear somewhat suspect.

As an initial response to legal mandates, then, organizations elaborate their structures to demonstrate attention to law. But there are important differences among organizations both in the timing of their responses and in the types of symbolic structures that they create. Organizations most sensitive to their normative environments are the innovators of structural responses and are more likely to create high-investment EEO/AA structures. As societal approval for EEO/AA in employment increases, and as it becomes apparent that EEO/AA structures have some advantages for the governance of employees, these structures help to erode managerial resistance to EEO/AA law. This engenders a second wave of structural elaboration, which is more universal but involves lower-investment EEO/AA structures.

CONCLUSION

This research suggests that, where legal ambiguity, procedural constraints, and weak enforcement mechanisms leave the meaning of compliance open to organizational construction, organizations that are subject to normative pressure from their environment elaborate their formal structures to create visible symbols of their attention to law. Structural elaboration helps to alleviate the conflict between legal norms and managerial interests by helping organizations to secure legitimacy as well as more tangible environmental resources while at the same time allowing administrators to preserve at least some managerial discretion.

Organizations' structural responses to law mediate the impact of law on society by helping to construct the meaning of compliance in a way that accommodates managerial interests. At the level of individual organizations, the construction of compliance becomes a function of internal organizational politics tempered by industry norms and the standards of professional personnel and affirmative action administrators (Edelman et

---

[33] The flag for imputed size is negative and significant at the .1 level (two-tailed test), suggesting that organizations that did not report size before 1984 are somewhat less likely to have rules than organizations that did report size before 1984.

al. 1991). At a broader level, organizations' structural responses to law help to shape legal and societal expectations about what constitutes compliance and good faith efforts to comply. When organizations claim that, by creating EEO/AA structures, they have eliminated discriminatory practices, they force courts, lawmakers, and society to struggle with the question of what constitutes compliance. Courts, for the most part, only legitimate or delegitimate forms of compliance that organizations devise. But it is important to keep in mind that most organizations' constructions of compliance are never examined in court. Thus organizations' collective response to law becomes the de facto construction of compliance; it is shaped only at the margins by formal legal institutions.

The institutionalization process, moreover, appears to render EEO/AA structures somewhat immune from changes in the political environment. It is striking that during the 1980s, 75 EEO/AA rules (31%) and 11 EEO/AA offices (17%) were created, even though the Reagan and Bush administrations significantly reduced pressure to comply with EEO/AA mandates. The continued creation of EEO/AA structures suggests that, over time, pressure shifts from the legal realm to the societal and organizational realms. As EEO/AA structures become institutionalized responses to law, personnel and affirmative action professionals are likely to institute these structures because of their apparent rationality; thus the waning political support has little immediate effect. And as attention to EEO/AA becomes more widespread, local minority and female communities may become more likely to demand change; in some cases, a new affirmative action officer may help to mobilize community or employee demands for change (Edleman et al. 1991).

Internal and societal pressures also mean that waning political support is unlikely to result in significant dismantling of EEO/AA structures. In my sample, only three offices were dismantled during the entire period of observation. Given the symbolic value of EEO/AA structures, to dismantle an EEO/AA office or rule would appear to be a flagrant sign of disdain for civil rights; it would be likely to provoke protest even in a depoliticized environment. I would expect that a reduction in political pressure might lead organizations to cut the budget or staff of an EEO/AA office rather than to dismantle it.

The elaboration of organizational structure is only the initial stage of the compliance process. While I have emphasized the symbolic value of EEO/AA structures, it remains uncertain at this point whether these structures act as a stepping stone toward the achievement of EEO/AA ideals or whether they exist as mere window dressing. Critical theories of law, which emphasize the fallacy of formal rights, would suggest that symbolic structures are probably the end of employers' response to law.

But traditional sociology of law, which holds that much of law's effect occurs in its shadow, suggests that, once in place, EEO/AA structures may produce or bolster internal constituencies that help to institutionalize EEO/AA goals.

Whether EEO/AA structures bring about more changes is an empirical question that is outside the scope of this article, but one observation is in order. Leonard (1986) reports that the greatest gains in the work-force shares of minorities and women occurred in the latter part of the 1970s and that these changes are most notable among federal contractor organizations. The "bottom-line" improvements in the work-force positions of minorities and women, then, appear to follow (with a lag of several years) the surge in EEO/AA office formation that begins around 1970 and are consistent with the higher rates of office formation by contractors before that time. This apparent relation suggests the possibility of a "shadow of the law" effect. But it is clear from this research that EEO/AA law (and other law with similiar characteristics) invites forms of organizational compliance that do not *guarantee* substantive results.

APPENDIX A

Survey of EEO/AA Practices

Sample

The survey of EEO/AA practices was administered to a national stratified probability sample of 350 organizations. The sample consists of 248 business organizations, 51 colleges (or universities), and 51 government agencies (10 federal, 21 state, and 20 county or municipal). The three types of organizations were drawn from separate sampling frames.

The business sample was drawn from the Large Company Data Base of Trinet, Inc. (Parsippany, N.J.), which consists of more than 215,000 listings of U.S. companies with 20 or more employees. In this data base, "company" denotes branches and subsidiaries as well as main offices and parent companies. Trinet updates and verifies its company files constantly, and their information is drawn from a variety of sources, including several proprietary business information data bases and data on new companies gathered in a telemarketing center used for business-to-business applications. For this survey, a sample was selected from the universe of U.S. companies with 100 or more employees in the 48 contiguous states. Five types of businesses were excluded from the sampling frame based upon their Standard Industrial Classification (SIC) codes: agricultural production, general building contractors, heavy construction

contractors, special trade contractors, and educational services.[34] The sampling frame of companies defined by the above criteria was stratified into two groups: one from the 10 states in which EEO records are available to the public (for use in later matching) and the remaining contiguous states. The sampling frame was further stratified by number of employees: 100–499 employees and 500 or more employees. And, within each of the four groups, the sampling frame was stratified by region and SIC code classification. A random selection of companies was then drawn within each stratum.

The college and university sample was drawn from a sampling frame of institutions of higher education accredited by the U.S. Department of Education. This frame is itself based on a 1986 survey of accredited institutions in the U.S., which had a response rate of 97.4%. Total enrollment (undergraduate, graduate, full- and part-time) was used as a proxy measure for size of staff for stratification purposes. The sampling frame was stratified by size of enrollment: the colleges that accounted for the top 30% of all enrollment (3,076 or more students) were designated as large; the remaining 70% were designated as small. Equal numbers were drawn from each size stratum. Within each size stratum, the sampling frame was further stratified by region and status of the institution as public or private.

The government sample was drawn from the Carroll Publishing *Directories of Government* (Washington, D.C.). Separate samples were drawn for federal/federal regional, state, and county/municipal agencies. The Carroll Publishing directories are a widely recognized source for government contacts at all levels of government, and are updated and verified constantly by a research staff in Washington, D.C. Defense industry agencies were excluded from the sampling frame.

Survey

The survey, which was conducted for me by the Gallup Organization, consisted of a 15–25 minute phone interview followed by a 10-page mail survey.[35] The telephone interview focused on general information about the organization relating to EEO/AA practices as well as general organizational characteristics. The mail survey consisted of more specific infor-

---

[34] Agricultural production, general building contractors, and heavy construction contractors were excluded because of the different nature of the employment relation in those industries and, in the case of construction contractors, because of differences in the laws regulating the employment relation. Educational services was excluded because educational organizations were sampled separately.

[35] With the exception of the dates of rule creation for some organizations, all data used in this article is based on data from the telephone survey.

mation regarding the nature of EEO/AA problems and complaints in organizations.

Each of the sampled locations was contacted by telephone, and the appropriate respondent was identified when the caller described the survey briefly and asked to speak with "the person with the most responsibility for affirmative action and equal employment opportunity policies and programs." This procedure was repeated until the respondent said that he or she was the person who could best answer the survey. The respondent in most cases was the affirmative action officer or personnel director.

Response Rates

Within each sampling frame and stratum, a sufficient number of organizations were drawn to obtain the desired distribution of completed interviews. The response rates for the business, college, and government samples are calculated separately because of the differences in sampling frames. The response rate for each of the telephone samples was calculated as the product of three component rates: the contact rate (the proportion of organizations in the sampling frame with working phone numbers that were contacted); the cooperation rate (the proportion of contacted organizations that cooperated with screening); and the completion rate (the proportion of organizations that cooperated with the screening in which interviews were completed). The response rates were 56.0% for the business sample, 49.5% for the college sample, and 44.7% for the government sample.

The mail survey was sent to respondents after completion of the telephone survey. The mail response rates, which were calculated simply as the proportion of surveys sent that were completed and returned, were 57.3% for the business sample, 62.7% for the college sample, and 68.6% for the government sample.

In cleaning the data, I dropped four of the completed interviews from the survey (two because most of the data were missing and two because the wrong type of organization had been selected). Thus the final sample consists of 248 business organizations, 50 colleges, and 48 government agencies (8 federal, 20 state, and 20 local).

Data

The data consist of 275 variables, including information on organizational characteristics (e.g., size, industry), organizations' administrative and contractual linkages to federal and state governments, the structure of personnel administration, and detailed questions about the structure of EEO/AA practices. The questions on EEO/AA practices include infor-

mation on changes in structure to handle EEO/AA requirements (such as the establishment of procedures to handle EEO/AA disputes, special EEO/AA offices, rules, counselors, affirmative action plans, and recruitment and training programs), how EEO/AA offices and EEO/AA activities are staffed, the autonomy and authority of EEO/AA offices, the role of lawyers in EEO/AA activities, the number of EEO/AA complaints filed internally and externally in the year 1986, and the timing of any EEO/AA lawsuits filed against organizations. Much of the data is event-history data (it includes dates that structures were created).

REFERENCES

Allison, Paul D 1984. *Event History Analysis: Regression for Longitudinal Event Data*. Beverly Hills, Calif.. Sage.

Ashenfelter, Orley, and James Heckman. 1976. "Measuring the Effect of an Anti-Discrimination Program." Pp. 46–48 in *Evaluating the Labor-Market Effects of Social Programs*, edited by Orley Ashenfelter and James Blum. Princeton, N.J.: Princeton University Press.

Baron, James N., Frank R. Dobbin, and P. Devereaux Jennings. 1986. "War and Peace: The Evolution of Modern Personnel Administration in U.S. Industry." *American Journal of Sociology* 92:350–83.

Beller, Andrea. 1982. "The Impact of Equal Opportunity Policy on Sex Differentials in Earnings and Occupations." *American Economic Review* 72:171–75.

Belton, Robert. 1981. "Discrimination and Affirmative Action. An Analysis of Competing Theories of Equality and *Weber*." *North Carolina Law Review* 59:531–98.

Blumrosen, Alfred W. 1965 "Anti-Discrimination Laws in Action in New Jersey. A Law-Sociology Study." *Rutgers Law Review* 19:187–287.

———. 1972. "Strangers in Paradise: *Griggs v. Duke Power Co.* and the Concept of Employment Discrimination." *Michigan Law Review* 71:59–110.

Bound, John, and Richard B. Freeman. 1989. "Black Progress: Erosion of the Post-1965 Gains in the 1980s?" Pp. 32–49 in *The Question of Discrimination: Racial Inequality in the U.S. Labor Market*, edited by Steven Shulman and William Darity, Jr. Middletown, Conn.: Wesleyan University Press.

Bullock, Charles S., III, and Charles M. Lamb, eds. 1984. *Implementation of Civil Rights Policy*. Monterey, Calif.: Brooks/Cole.

Bumiller, Kristin. 1987. "Victims in the Shadow of the Law: A Critique of the Model of Legal Protection." *Signs* 12:421–34.

———. 1988. *The Civil Rights Society*. Baltimore: Johns Hopkins University Press.

Burk, James. 1988. *Values in the Marketplace: The American Stock Market under Federal Securities Law*. New York: Walter de Gruyer.

Burstein, Paul. 1990. "Intergroup Conflict, Law, and the Concept of Labor Market Discrimination." *Sociological Forum* 5:459–76.

Burstein, Paul, and Kathleen Monaghan. 1986 "Equal Employment Opportunity and the Mobilization of Law." *Law and Society Review* 20:355–88.

Butler, Richard, and James J. Heckman. 1977. "The Government's Impact on the Labor Market Status of Black Americans: A Critical Review." Pp. 235–81 in *Equal Rights and Industrial Relations*, edited by Leonard J. Hausman, Orley Ashenfelter, Bayard Rustin, Richard F. Schubert, and Donald Slaiman. Champaign, Ill : Iowa Graphics.

Carlin, Jerome E., and Jan Howard. 1965. "Legal Representation and Class Justice." *UCLA Law Review* 12:381–437

Carroll, Glenn R. 1983. "Dynamic Analysis of Discrete Dependent Variables: A Didactic Essay." *Quality and Quantity* 17·425–60.

Chamberlain, Neil W., Donald E. Cullen, and David Lewin. 1980. *The Labor Sector.* New York: McGraw-Hill.

Clune,. William H. 1983. "A Political Model of Implementation and the Implications of the Model for Public Policy, Research, and the Changing Role of Lawyers." *Iowa Law Review* 69:47–125.

Coleman, James S. 1984. "Introducing Social Structure into Economic Analysis." *American Economic Review* 74 (2): 84–88.

Crenshaw, Kim. 1988. "Race, Reform and Retrenchment: Transformation and Legitimation in Antidiscrimination Law." *Harvard Law Review* 101 1131–1387.

Cyert, Richard M., and James G. March. 1963. *A Behavioral Theory of the Firm.* Englewood Cliffs, N.J. Prentice-Hall

DiMaggio, Paul J., and Walter W. Powell. 1983. "The Iron Cage Revisited. Institutional Isomorphism and Collective Rationality in Organizational Fields." *American Sociological Review* 48:147–60.

Eberts, Randall W., and Joe A. Stone. 1985 "Male-Female Differences in Promotions: EEO in Public Education " *Journal of Human Resources* 20.504–21

Edelman, Lauren B. 1990. "Legal Environments and Organizational Governance. The Expansion of Due Process in the American Workplace." *American Journal of Sociology* 95:1401–40.

Edelman, Lauren B., Steven E. Abraham, and Howard S. Erlanger. 1992. "Professional Construction of Law The Inflated Threat of Wrongful Discharge." *Law and Society Review,* in press.

Edelman, Lauren B., Stephen Petterson, Elizabeth Chambliss, and Howard S. Erlanger. 1991. "Legal Ambiguity and the Politics of Compliance: Affirmative Action Officers' Dilemma." *Law and Policy* (Spring), in press.

Eisenberg, Theodore. 1990. "Testing the Selection Effect: A New Theoretical Framework with Empirical Tests." *Journal of Legal Studies* 19.337–58.

Felstiner, William L. F., Richard L. Abel, and Austin Sarat. 1981. "The Emergence and Transformation of Disputes: Naming, Blaming, Claiming . . . " *Law and Society Review* 15:631–54.

Fiss, Owen M. 1974. "The Fate of an Idea Whose Time Has Come Antidiscrimination Law in the Second Decade after *Brown v Board of Education.*" *University of Chicago Law Review* 41.742–73.

Freeman, Alan. 1982. "Legitimating Racial Discrimination through Anti-Discrimination Law. A Critical Review of Supreme Court Doctrine." Pp. 210–35 in *Marxism and Law,* edited by Piers Beirne and Richard Quinney. New York: Wiley

Freeman, Richard B., and James L. Medoff. 1984. *What Do Unions Do?* New York: Basic Books.

Froehlich, Herbert P , and Dennis A. Hawver 1974. "Compliance Spinoff: Better Personnel Systems." *Personnel* (January–February):62–68.

Galanter, Marc. 1974. "Why the 'Haves' Come Out Ahead: Speculations on the Limits of Legal Change " *Law and Society Review* 9:95–160.

Garris, Steve, and Ann Black. 1974. "Revising Personnel Management Procedures." *Personnel* (November–December):50–58.

Hawkins, Keith. 1984. *Environment and Enforcement.* Oxford: Clarendon Press.

Heckman, James J., and Brook S. Payner. 1989. "Determining the Impact of Federal Antidiscrimination Policy on the Economic Status of Blacks: A Study of South Carolina " National Bureau of Economic Research Working Paper no. 2854. Cambridge, Mass.

Heckman, James J., and J Hoult Verkerke. 1990. "Racial Disparity and Employment Discrimination Law· An Economic Perspective." *Yale Law and Policy Review* 8:276–99.

Hill, Herbert. 1975. "Black Labor, the NLRB, and the Developing Law of Equal Employment Opportunity." *Labor Law Journal* 26:207–23.

House Committee on Education and Labor. 1987. *A Report on the Investigation of the Civil Rights Enforcement Activities of the Office of Federal Contract Compliance Programs, U.S. Department of Labor.* 100th Congress, 1st sess. Serial no. 100-R.

Jacoby, Sanford M 1985 *Employing Bureaucracy: Managers, Unions, and the Transformation of Work in American Industry, 1900–1945.* New York Columbia University Press.

Jaffe, Louis L. 1954. "The Effective Limits of the Administrative Process A Reevaluation." *Harvard Law Review* 67:1105–35.

Jones, James E. 1970. "Federal Contact Compliance in Phase II—the Dawning of the Age of Enforcement of Equal Employment Obligations." *Georgia Law Review* 4:756–78.

———. 1982. "Twenty-One Years of Affirmative Action: The Maturation of the Administrative Enforcement Process under the Executive Order 11,246 as Amended " *Chicago Kent Law Review* 59.67–122.

———. 1985a. "The Genesis and Present Status of Affirmative Action in Employment: Economic, Legal, and Political Realities." *Iowa Law Review* 70:901–44.

———. 1985b. "Some Reflections on Title VII of the Civil Rights Act of 1964 at Twenty " *Mercer Law Review* 36:813–53.

Kaplan, E. L., and P. Meier 1958. "Nonparametric Estimation from Incomplete Observations." *Journal of the American Statistical Association* 53:457–81

Leonard, Jonathan S. 1984a "Antidiscrimination or Reverse Discrimination. The Impact of Changing Demographics, Title VII, and Affirmative Action on Productivity." *Journal of Human Resources* 19:145–74.

———. 1984b. "The Impact of Affirmative Action on Employment." *Journal of Labor Economics* 2:439–63.

——— 1985. "What Promises Are Worth: The Impact of Affirmative Action Goals." *Journal of Human Resources* 20:3–20.

———. 1986. "The Effectiveness of Equal Employment Law and Affirmative Action Regulation." *Research in Labor Economics* 8:319–50.

Loevi, Francis J., Jr. 1973. "The Union Role in Federal EEO Programs." *Public Personnel Management* 2:162–66.

Marino, Kenneth E. 1980. "Conducting an Internal Compliance Review of Affirmative Action." *Personnel* (March-April):24–34

Meyer, John W., and Brian Rowen. 1977. "Institutionalized Organizations· Formal Structure as Myth and Ceremony."*American Journal of Sociology* 83·340–63.

Meyer, John W., and W. Richard Scott, eds 1983. *Organizational Environments: Ritual and Rationality* Beverly Hills, Calif.: Sage.

Miller, Richard E., and Austin Sarat 1981. "Grievances, Claims and Disputes: Assessing the Adversary Culture." *Law and Society Review* 15:525–66.

Nathan, Richard P. 1969. *Jobs & Civil Rights. The Role of the Federal Government in Promoting Equal Opportunity in Employment and Training.* Clearinghouse Publication no. 16. Prepared for the U.S. Commission on Civil Rights by the Brookings Institution. Washington, D.C. Government Printing Office.

Newman, Winn, and Carole W. Wilson. 1981. "The Union Role in Affirmative Action." *Labor Law Journal* 32:323–42.

Schultz, Vicki. 1990. "Telling Stories about Women and Work: Judicial Interpretations of Sex Segregation in the Workplace in Title VII Cases Raising the 'Lack of Interest' Argument." *Harvard Law Review* 103·1750–1843.

Scott, W. Richard. 1983. "The Organization of Environments: Network, Cultural and Historical Elements." Pp. 129–53 in *Organizational Environments: Ritual and Rationality,* edited by John W. Meyer and W. Richard Scott Beverly Hills, Calif.: Sage.

Scott, W Richard, and John W. Meyer. 1983. "The Organization of Societal Sectors." Pp. 129–54 in *Organizational Environments: Ritual and Rationality,* edited by John W. Meyer and W. Richard Scott. Beverly Hills, Calif.. Sage.

Selznick, Philip P. 1969. *Law, Society, and Industrial Justice.* New Brunswick, N.J.: Transaction

Smith, Arthur B., Jr. 1978. *Employment Discrimination Law.* Indianapolis, Ind · Bobbs-Merrill.

Stone, Christopher. 1975. *Where the Law Ends: The Social Control of Corporate Behavior* New York. Harper & Row

Thorp, Cary D., Jr 1973. "Fair Employment Practices: The Compliance Jungle." *Personnel Journal* 52.642–49.

Tushnet, Mark. 1984. "An Essay on Rights." *Texas Law Review* 62:1363–1403.

U.S Commission on Civil Rights (USCCR). 1971. *The Federal Civil Rights Enforcement Effort: One Year Later.* Clearinghouse Publication no. 34. Washington, D C : Government Printing Office.

———. 1986. *The Economic Progress of Black Men in America.* Clearinghouse Publication no 91. Washington, D.C.: Government Printing Office.

Vaughan, Diane 1982 "Transaction Systems and Unlawful Organizational Behavior." *Social Problems* 29:373

Weick, Karl E. 1976. "Educational Organizations as Loosely Coupled Systems."*Administrative Science Quarterly* 21:1–19

Whalen, Garry M., and Richard S. Rubin. 1977. "Labor Relations and Affirmative Action: A Tug-of-War." *Public Personnel Management* 6 149–55.

Wirt, Frederick. 1970. *The Politics of Southern Equality: Law and Social Change in a Mississippi County.* Chicago. Aldine

Wood, Norman J. 1975. "Equal Employment Opportunity and Seniority: Rights in Conflict." *Labor Law Review* 26:345–49.

Statutes

Equal Pay Act of 1963, 29 U.S.C. §206(d).
Title VII of the 1964 Civil Rights Act, 42 U.S.C. §2000e
Age Discrimination in Employment Act of 1967, 29 U.S.C. §621.
Equal Employment Opportunity Act of 1972, P.L. 92–261, 86 Stat. 103.
Rehabilitation Act of 1973, 29 U S C. §701 et seq.
Vietnam Era Veterans' Readjustment Assistance Act of 1972, 38 U.S.C. §2011.
Vietnam Era Veterans' Readjustment Assistance Act of 1974, 38 U S C. §2012.

Executive Orders

Executive Order 8802, 3 C.F.R. 957 (1938–1943 Comp.).
Executive Order 10925, 3 C F.R. 448 (1961).
Executive Order 11246, 3 C.F.R. 339 (1964–1965 Comp.).
Executive Order 11375, 32 Fed Reg, 14303 (1967).
Executive Order 12086, 43 Fed Reg. 46501 (1978).

Administrative Rules

EEOC Affirmative Action Guidelines of 1979, 29 C.F.R. §1608 (1980).
OFCC Rules and Regulations (includes Order 4) 35 Fed. Reg. 2586 (1970) (codified in 41 C.F.R. Part 60-2).
OFCC Rules and Regulations (includes Revised Order 4) 36 Fed. Reg. 23452 (1971) (codified in 41 C.F.R. Part 60-2).

Cases

EEOC v Sears, Roebuck & Co. 628 F. Supp. 1264 (N.D. Ill. 1986), aff'd 839 F.2d 302 (7th Cir. 1988).
Griggs v. Duke Power Co. 401 U.S. 424, 91 S Ct. 849, 28 L.Ed. 2d 158 (1971)
Lorance v. AT&T Technologies, Inc. 490 U.S. 900, 109 S Ct. 2261, 104 L.Ed. 2d 961 (1989).
Mazus v. Dept. of Transportation, Commonwealth of Pennsylvania, 629 F.2d 870 (3d Cir. 1980), cert. denied, 449 U.S. 1126, 101 S.Ct. 945, 67 L.Ed. 2d 113 (1981).
Movement for Opportunity and Equality v. Detroit Diesel Allison Division of General Motors Corp. 18 FEP 557 (S D. Ind. 1978), aff'd 622 F.2d 1235 (7th Cir. 1980).
Price Waterhouse v. Hopkins 490 U.S. 228, 109 S.Ct. 1775, 104 L.Ed. 2d 268 (1989).
Ste. Marie v. Eastern Railroad Association 650 F.2d 395 (2d Cir. 1981).
Wards Cove Packing, Inc. v. Atonio, 490 U.S. 642, 109 S.Ct. 2115, 104 L.Ed (1989)

Reflected Appraisals, Parental Labeling, and Delinquency: Specifying a Symbolic Interactionist Theory[1]

Ross L. Matsueda
University of Wisconsin—Madison

This article draws on behavioral principles of George Herbert Mead and other symbolic interactionists to specify a theory of the self to explain delinquent behavior. The theoretical framework builds on Mead's analysis of the social act, symbolic interactionists' specification of the self as a reflection of appraisals made by significant others, and labeling theorists' notions of dramatization of evil, deviance amplification, and secondary deviance. This integrated framework is tested with a causal model of the causes and consequences of reflected appraisals and delinquent behavior. The analysis provides general support for the theory. Reflected appraisals of self are substantially affected by parental appraisals and prior delinquency; future delinquency is substantially affected by reflected appraisals of self as a rule violator; and reflected appraisals mediate much of the effects on delinquency of parental appraisals, prior delinquency, and structural variables.

An important question in the study of social control involves the mechanisms by which informal groups control the behavior of members. Much research on informal controls and delinquent behavior has examined rela-

tionships between parental socialization, self-concepts, and delinquency. As Wells and Rankin (1983) put it, self-concepts should be an important mediating factor in delinquency, intervening between parental socialization and delinquent behavior. Accordingly, researchers have produced a voluminous literature that investigates the relationship between self-concepts and delinquency. The results of that research have been disappointing. When conceptualized as global self-esteem or self-rejection, self-concepts appear to have modest or inconsistent effects on delinquent behavior. These results suggest the need for considering alternative conceptualizations of the self and its role in the process of social control.

In this article I will draw on the writings of George Herbert Mead (1934) and the school of symbolic interactionism to conceptualize the self as being rooted in social interaction, comprising multiple dimensions, and providing a crucial link between self-control and social control. I will draw on theories of labeling and reference groups to specify the broader determinants of the self and argue that delinquency is in part determined by one's appraisals of self from the standpoint of others.

THEORY AND RESEARCH ON DELINQUENCY AND SELF-CONCEPTS

Most research and theory about the self and delinquency has focused on global self-esteem. In perhaps the best theoretical statement on self-esteem, Rosenberg (1979) argues that the formation of global self-esteem entails three mechanisms: reflected appraisals, social comparison, and self-attribution (see also Rosenberg and Simmons 1972). Through the process of reflected appraisals, individuals form self-conceptions on the basis of their perceptions of others' attitudes toward them. Through the process of social comparisons, people make judgments about themselves, in part by comparing themselves with others (Festinger 1954). And through the process of self-attribution, individuals draw conclusions about their dispositions, motives, and self-esteem on the basis of their observations of their own overt behavior (Bem 1972). These mechanisms imply that a strong motive for delinquent behavior, as well as for other forms of behavior, is the acquisition and maintenance of high self-esteem. Adolescents may turn to delinquency to enhance their self-esteem (Kaplan 1975) or to overcome feelings of self-rejection (Kaplan 1980). Thus, positive self-esteem may insulate one from delinquency (Reckless, Dinitz, and Murray 1956).

Recent empirical research on self-esteem and delinquency, which capitalizes on longitudinal data, has produced equivocal results. Several sets

of analyses of the Youth in Transition data set (Bachman, O'Malley, and Johnston 1978) reveal conflicting results: some have found support for the self-enhancement principle (self-esteem affects delinquency; see Rosenberg and Rosenberg 1978; Rosenberg, Schooler, and Schoenbach 1989), while others have not (Bynner, O'Malley, and Bachman 1981); and some have found support for the reflected appraisals principle (delinquency affects self-esteem; see Wells and Rankin 1983; McCarthy and Hoge 1984), while others have not (Rosenberg and Rosenberg 1978). Furthermore, in a series of analyses, Kaplan and his colleagues have found consistent support for the self-enhancement principle. They found that, net of three variables (prior deviance, deviant peers, and disposition to deviance), prior self-rejection exerts a small but significant effect on future deviance (e.g., Kaplan, Johnson, and Bailey 1987).

Given that global self-esteem appears to have modest effects on delinquency, it may be fruitful to examine other conceptualizations of the self and self-control (Wells and Rankin 1983, p. 20). A promising framework for analyzing delinquency and the self is symbolic interactionism, a perspective that includes an explicit theory of the self and social control. From an interactionist perspective, global self-esteem is only one element of a multifaceted self and may not be the most important determinant of delinquent behavior (Wells 1978). A more important determinant may be the specific content or meaning of the self that is relevant to delinquent behavior (Schwartz and Stryker 1970), such as evaluations of the self as a delinquent versus evaluations as a conformist. Furthermore, the critical locus of social control may be the process of role-taking and forming the self as an object with a specific set of meanings. Thus, a useful approach would examine the self (as delinquent or conformist) as a reflection of the appraisals of others—not as a principle governing the formation of global self-esteem, as specified by Rosenberg (1979)—but rather as one component of a symbolic interactionist principle of social control.

An early attempt to examine deviance from an interactionist perspective was conducted by Schwartz and Stryker (1970), who hypothesized that boys labeled as "bad" by teachers should be more likely than boys labeled "good" (1) to have poor and uncertain self-concepts; (2) to exclude members of conventional institutions (teachers) as significant others; and (3) to have more difficulty with masculine identities. Their results provided mixed support for these hypotheses. While only partially successful, this study remains the only major empirical study of an interactionist approach to deviance (Stryker and Craft 1982). I will attempt to build on this research here by developing an explicit interactionist theory of the self and delinquency and subjecting the theory to empirical test.

SELF-CONTROL AS SOCIAL CONTROL· A CONCEPTION OF SELF BASED ON MEAD

The perspective of symbolic interactionism presupposes that social order is the product of an ongoing process of social interaction and communication. Of central importance is the process by which shared meanings, behavioral expectations, and reflected appraisals are built up in interaction and applied to behavior. These shared meanings attach to positions in society and thus link individual conduct to the organization of groups and to social structure. Social structure—the patterned regularities in society—is an ongoing process, built up by social interactions; moreover, social structure in turn constrains the form and direction of these interactions by structuring communication patterns, interests, and opportunities (Stryker 1980). The specific mechanism linking interaction and social structure is role-taking.[2]

Role-Taking and Delinquency

To analyze interaction, symbolic interactionists define the unit of analysis as the transaction, which consists of an interaction between two or more individuals. Within transactions, the important mechanism by which interactants influence each other is role-taking, which consists of projecting oneself into the role of other persons and appraising, from their standpoint, the situation, oneself in the situation, and possible lines of action. With regard to delinquency, individuals confronted with delinquent behavior as a possible line of action take each other's roles through verbal and nonverbal communication, fitting their lines of action together into joint delinquent behavior (Mead 1934; Blumer 1969).

The transaction is built up through this dynamic process of reciprocal role-taking: one person initiates action—say, an unlawful act—a second takes the role of the other and responds, then the first person reacts to the response, and so on, until the jointly developed goal is reached, a new

[2] This perspective emphasizes the study of patterns of behavior and meanings that remain relatively stable across a delimited set of situations. Somewhat stable meanings can be examined with quantitative survey data. This view is consistent with the methodological thrust of "structural symbolic interactionism," which stresses the structure of role relationships that generate stable meanings and behaviors (Stryker 1980, McCall and Simmons 1978). It is less consistent with the methodological recommendations of Blumer's (1969) symbolic interactionism, which stresses the negotiated, interpreted, and constructed nature of meaning, and eschews "variable analysis" for the study of human experiences that give rise to meanings. Note, however, that it is consistent with Blumer's (1969, p. 139) conclusion that "in the area of interpretative life, variable analysis can be an effective means of unearthing stabilized patterns of interpretation which are not likely to be detected through the direct study of the experience of people "

goal is substituted, or the transaction simply fades. Through reciprocal role-taking, or a conversation of gestures, consensus over situational goals and the appropriate means for attaining those goals is constructed, individual lines of action are coordinated, and there is concerted action toward achieving the goal (Blumer 1969). Thus, the initiated delinquent act of one youth might elicit a negative response from another youth, causing the group to search for another, more suitable alternative. Whether or not a goal is achieved using unlawful means is determined by each individual's contribution to the direction of the transaction; those contributions, in turn, are determined by the individual's prior life experience or biography (Hewitt 1988).

Early in the socialization process, individuals engage in a serial process of taking the role of specific significant others who are present in the interaction. Later in the socialization process, individuals learn to take the role of the entire group or "generalized other," which includes the norms, rules, and expectations governing various positions and roles of a group, community, or society. Here, individuals learn to relate the activities and expectations of their roles to the activities and expectations of other roles within an organized system (Mead 1934, pp. 152–64). This form of taking the role of an organized and abstract group appears in more institutionalized settings and constitutes the most effective form of social control, since the organized institutions and norms enter individual behavior.

Role-taking also provides a framework for an interactionist theory of cognition. Cognitive processes arise in problematic situations, in which a line of action (impulse) is temporarily blocked by physical objects in the situation, by verbal responses of others, or by subjective reactions such as repugnance, shame, and fear (Shibutani 1961). The blocked impulse is transformed into a self-image (the self as an object or the "me"), consisting of alternative lines of action, anticipated reactions of others, and, most significantly, a view of self from the standpoint of others. The line of action is then reacted to by another impulse (the "I"), which either reacts positively and follows the line of action into overt behavior or reacts negatively, blocking the impulse to act and eliciting another self-image. This cognitive process continues until the problem is solved or the transaction ends. Thus, cognition is identical, in form and content, to role-taking between interactants, except that it occurs in the mind in an imaginative rehearsal between the "I" and the "me" (Mead 1934).

Moreover, similar situations will call out similar "me's"—the self formed as an object from the standpoint of others. Therefore, a stable self arises because the self-images ("me's") called up in a situation, to which the "I" will react, will resemble previous "I's" and "me's" from similar past situations. This stable set of self-images is multidimensional,

containing an organized set of stable meanings about oneself from the standpoint of others. Mead (1934, p. 142) termed this self "multiple personality" to emphasize that it is a reflection of the organized social process; McCall and Simmons (1978) and Stryker (1980) conceptualized it as "role-identities" to emphasize that it corresponds to the many social roles one plays; and Kinch (1963) conceived it as "reflected appraisals" to emphasize that it is a reflection of appraisals made by significant others. With regard to delinquency, the important element of the self formed as an object is the specific meaning or content of the self with respect to delinquency. Those who see themselves (from the standpoint of others) as persons who engage in delinquent behavior in certain situations are more likely to engage in delinquency. Thus, if the self as a delinquent is an important dimension of the self for individuals, such that it endures across situations, it should predict individuals' delinquent behavior.

Most behavior, particularly in highly institutionalized and routinized transactions, occurs in nonproblematic situations and results from nonreflective habitual behavior, based on the way in which previous problematic situations were resolved. When a problematic situation is repeatedly encountered, it becomes less problematic, as one learns to resolve it proficiently. Eventually, the situation becomes nonproblematic and the behavior habitual. This implies that over time, delinquent behavior will become increasingly stable, so long as one encounters similar situations. Of course, behavior will not be completely stable because situations are in part selected through cognitive processes, and the response of the "I" is not completely determined by the "me."[3]

The process by which role-taking can lead to delinquent behavior is illustrated by four classic studies of delinquency. Briar and Piliavin (1965) found that boys freed from commitments to conventional lines of action are often incited into delinquency by "situationally induced motives," which are verbal motives presented by other boys. Free from considering the reactions of conventional others, these boys can take the role of each other, present delinquent motives, and jointly adopt a delinquent line of

[3] The use of the social transaction as the unit of analysis is a convenient abstraction from the ongoing social process used for analytical purposes. Mead (1938) specified the social act as beginning with the stage of impulse and ending with consummation, with perception and manipulation intervening. Note, however, that the consummation stage of one transaction may be the impulse stage for another. This implies that individuals place themselves in certain situations through habit or through cognitive processes arising from a previous problematic situation. Unlike radical behaviorism, symbolic interactionism does not subscribe to what Wrong (1961) termed "the oversocialized conception of man" (Stryker 1980). Individuals are not viewed as completely determined beings, passively conforming to expectations or reinforcement contingencies but instead are seen as active beings in part constrained by social organization (through the "me") and in part creating that organization (through the "I").

action. Short and Strodtbeck (1965) noted that one's decision to join a gang fight often revolves around the risk of losing status within the gang. Gang members would take the role of the group, consider the group's negative reactions, and join in on the action for fear of losing status. Cohen (1955) argued that adolescent groups engage in a tentative probing conversation of gestures—a process best characterized as one of trial and error—and collectively innovate a new status hierarchy, a delinquent subculture. Finally, Gibbons (1971) claimed that as a result of group interactions, novel shades of norms and values emerge to influence the direction of joint behavior (Short 1974). Such processes, consistent with Smelser's (1963) "value added" and Turner's (1964) "emergent norm" approaches to collective behavior, show how a group controls the behavior of its members within a situation.

This discussion of role-taking implies four major features for a theory of the self and delinquent behavior. First, the self consists of an individual's perception of how others view him or her, and thus, is rooted in symbolic interaction. Second, the self as an object arises partly endogenously within situations, and partly exogenously from prior situational selves being carried over from previous experience. This results because self-images ("me's") called up in a situation will resemble previous "me's," while the "I" will respond in novel ways arising from the immediate situation.[4] Thus, we can speak of a set of patterned selves that is somewhat stable over time but varies across individuals. Third, the self as an object is a *process* determined by the self at a previous point in time and by prior behavior (resolutions of problematic situations). Fourth, delinquent behavior will result in part from the formation of habits and in part from stable perceptions of oneself from the standpoint of others. Through the latter process, delinquency is controlled by one's reference groups.

Role-Taking, Reference Groups, and Delinquency

Role-taking usually entails taking the role of members of one's reference group, which is a group that serves as a source of one's values, perspec-

___

[4] The "me" called up to solve a problematic situation will resemble previous "me's" because of stability in one's generalized others. That is, an individual will take the role of the generalized other, forming a "me," and consider alternative lines of action from the standpoint of that generalized other, which represents the organized structure of the group (Mead 1934, p. 199–201). Since one's generalized others remain somewhat stable across situations, there will be continuity in "me's" across situations. Furthermore, since one's generalized others are predicated on one's role commitments, the structure of the self that conditions cognition is organized by commitments (see Stryker 1980; McCall and Simmons 1978).

tives, and self-comparisons. Reference groups consist of individual significant others, such as parents, friends, and teachers, but also organized groups (generalized others) such as classmates, gangs, and families. In mass societies, members have multiple reference groups; which significant other or reference group is invoked *within a given situation* depends on many factors, the most important of which is the relevance of the group to the perceived problematic situation at hand. Moreover, those persons we care about, from whom we gain personal status, and who have helped form our self-image in the past, are most likely to be selected, since we want to maintain a favorable self-image in their eyes.

More broadly, an adolescent's multiple reference groups are determined by a complex set of individual variables, such as propinquity (Festinger 1954) and his or her perception that the group will provide a positive self-image (Hyman and Singer 1968). These individual determinants are structured by communication channels, which in turn are patterned by the larger social structure (Hewitt 1988, p. 125). I would expect that communication channels will be influenced by structural variables such as social class, family structure, residential area, and neighborhood structure, as well as individual characteristics such as age, race, sex, and cognitive ability. Thus, social structure should affect delinquency by structuring communication channels and reference groups, which in turn influence self-control—engaging in self-conscious reflective behavior. Therefore, self-control is social control because social structure enters behavior through role-taking, and because the self is constructed in a social process (Blumer 1969). We might term this "differential social control," since the direction of control—whether toward delinquency or toward conformity—differs by the problematic situation, the reference group, and the prior views of self by the individual (Glaser 1979).

Reflected Appraisals of Self and Delinquent Behavior

The foregoing discussion implies a specific conception of the self as a mechanism of social control. While the self as an object arises in problematic situations, we can also conceive of a self, in the form of consistent "me's," that is relatively stable across situations. Such a self, specified by Cooley (1922) as a "looking-glass self," and by Mead (1934) as the "self as an object," is a process consisting of three components: how others actually see one (others' actual appraisals); how one perceives the way others see one (reflected appraisals); and how one sees oneself (self-appraisals). Thus, one's self is in part a "reflected appraisal" of how significant others appraise one (Kinch 1963; Felson 1985).

Most empirical research on this conception of the self has followed

Miyamoto and Dornbush (1956) and examined the relationships among actual appraisals of significant others, reflected appraisals of significant others, and self-appraisals (Shrauger and Schoeneman 1979). This is diagrammed in the top third of figure 1. Felson (1985) argues that, according to symbolic interactionist theory, reflected appraisals of self are causes of self-appraisals and consequences of actual appraisals by others. Empirical research has found that actual appraisals have consistent but modest effects on reflected appraisals of self (Shrauger and Schoeneman 1979; Felson 1985, 1989). On the basis of these results, Felson (1980, 1989) suggests that (1) actual appraisals have only modest effects on reflected appraisals because of barriers to communication (peers do not always communicate their appraisals directly to their friends); (2) the effect is greater for appraisals that are socially defined in interaction (greater for attractiveness than for athletic performances); and (3) the relationship could be spurious owing to prior performance of the appraised behavior.
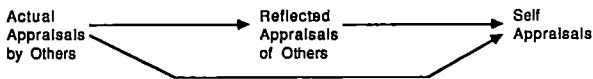
Felson's Model of Reflected Appraisals and Self Appraisals
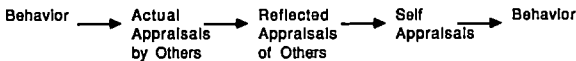
Kinch's Model of Reflected Appraisals and Behavior
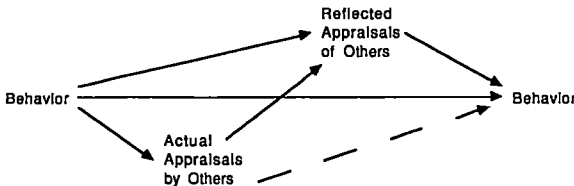
A Revised Model of Reflected Appraisals and Behavior

FIG. 1.—Alternative models of reflected appraisals

An interactionist conception of self as social control, however, does not imply a one-to-one correspondence between reflected appraisals and actual appraisals (Hewitt 1988, p. 129). Clearly, reflected appraisals are the result of *selective perception* of actual appraisals, which depends on the particular problematic situations that give rise to the reflected appraisals. Thus, reflected appraisals should be only partially a function of actual appraisals.

Research has also found that reflected appraisals have small effects on self-appraisals; however, these effects are larger than those of actual appraisals on self-appraisals (Shrauger and Schoeneman 1979; Felson 1985; Rosenberg and Simmons 1972). The inconsistent effect predicting self-appraisals may be due to social desirability effects or demand characteristics within the interview setting, which may be larger when evaluating oneself than when reporting how others evaluate one (Schrauger and Schoeneman 1979). Moreover, the theoretical framework outlined above suggests that the significant dimension of the self is reflected appraisals and not self-appraisals. If social control is exerted through role-taking, and the self that influences behavior is the object taken from the standpoint of others, it may be fruitful to examine the efficacy of reflected appraisals in explaining behavior.

There has been little research examining the relationship between the reflected appraisal process and actual behavior. Long ago, Kinch (1963) derived a theoretical model of reflected appraisals and behavior from symbolic interactionism, which posits a long causal chain (see the middle section of fig. 1). According to the model, initial behavior determines others' actual appraisals of a person, which in turn, lead to the person's reflected appraisals of self; reflected appraisals then determine self-appraisals, which in turn, lead to behavior. The model implies that, in the causal sequence explaining behavior, each antecedent variable in the model is entirely mediated by each subsequent variable. In light of the theoretical discussion above, I can derive a more plausible model of reflected appraisals and behavior.

The revised model, diagrammed in the bottom section of figure 1, follows Kinch by specifying that actual appraisals by others affect behavior only by affecting one's reflected appraisals of self. The alternative hypothesis, which contradicts symbolic interactionism, posits that actual appraisals influence behavior directly, regardless of reflected appraisals (indicated by a broken line in the bottom section of fig. 1). This could result if significant others are particularly proficient at appraising one and, therefore, predicting one's behavior or if other elements of the self besides reflected appraisals mediate actual appraisals. Moreover, the model diverges from Kinch's model in three ways. First, it deletes self-appraisals from the model, stipulating reflected appraisals of self as the

key variable for explaining behavior.[5] Second, it allows behavior to have a direct effect on subsequent behavior. This is consistent with our theoretical framework, which posits that institutionalized behavior, corresponding to Mead's (1934) nonreflective behavior and Dewey's (1922) habitual behavior, occurs in nonproblematic situations and is determined not by role-taking but by prior behavior. Third, it allows behavior to have a direct effect on reflected appraisals, since those appraisals are formed in part from previous behavioral solutions to problematic situations. Symbolic interactionism would predict that reflected appraisals are determined more by actual appraisals of others than by past behavior.

This last model can explain the relationships between parental appraisals, reflected appraisals, and delinquent behavior. It allows me to test three restrictions specified by Kinch (1963): (1) prior delinquency has no direct effect on later delinquency; (2) prior behavior has no direct effect on reflected appraisals; and (3) actual appraisals have no direct effect on future delinquency.[6] To link these social psychological mechanisms to broader determinants of delinquency, I turn to labeling theory.

THE PARENTAL CONTEXT OF CONTROL: LABELING AND REFLECTED APPRAISALS

Most etiological statements of labeling theory, particularly Tannenbaum's (1938) concept of the dramatization of evil, Lemert's (1951) concept of secondary deviance, and Mead's ([1918] 1964) concept of the hostile attitude of punitive justice, are rooted in the perspective of symbolic interactionism.[7] Therefore, we can draw on labeling theory to specify the broader social determinants of the reflected appraisal process (see Elliott, Ageton, and Canter 1979; Farrell and Swigert 1988).

Focusing primarily on the negative consequences of labeling an indi-

---

[5] In their examination of labeling theory and the I.Q.-delinquency debate, Menard and Morse (1984) found that perceived social labels (reflected appraisals) had large effects on delinquency.

[6] Note, however, that the revised model is unable to test Kinch's hypothesis that self-appraisals mediate the effects that reflected appraisals have on behavior. Such a test would require direct measures of self-appraisals.

[7] Here, I am referring specifically to the writings of Lemert (1951, 1972) and Tannenbaum (1938), which stress the causal consequences of "dramatization of evil" and "secondary deviance." What falls under the rubric of labeling theory typically includes the "societal reactions" perspective, which is not directly relevant to the present analysis. In some versions of societal reactions theory, advocates call for examining the process by which official processing leads to some persons being designated criminals, delinquents, or mentally ill (Kitsuse and Cicourel 1963). In other more extreme versions, proponents define deviance not as behavior but as a label conferred by a social audience (Becker 1963)

vidual as "deviant" or "delinquent," labeling theory argues that initial acts of delinquency are relatively harmless instances of primary deviance. From the standpoint of the child, such acts are defined as "play" or "mischief"; however, from the standpoint of the larger community, they are viewed as "evil" or as a "law violation." The community's response, which initially includes reactions of parents, teachers, and peers, and later encompasses reactions of the juvenile justice system, is to label the child as "bad" or "evil." The label, in turn, influences the self-image of the child, who comes to view him or herself as bad or delinquent, which in turn increases the likelihood of future deviance. Eventually, this spiraling labeling process can leave the youth in the hands of juvenile justice officials—cut off from conventional society, stigmatized by parents and teachers, and left with a delinquent self-image. Thus, a self-fulfilling prophecy is set up: through this process of deviance amplification, or secondary deviance, an otherwise conforming child may eventually respond to the initial labeling of harmless acts by confirming the delinquent label (Tannenbaum 1938). Mead (1964) argued that the hostile response of the criminal justice system, under the justification of deterrence or retribution, could operate to exacerbate rather than ameliorate the crime problem, perhaps creating a stable criminal class (Hagan and Palloni 1990).

A hallmark of labeling theory is the proposition that deviant labels are not randomly distributed across the social structure, but are instead more likely to apply to the powerless, the disadvantaged, and the poor. Because of existing stereotypes—which portray criminals as members of lower classes, minorities, urban dwellers, and young adults—individuals who belong to such groups are more likely than others to be labeled delinquent (Simmons 1965; Farrell and Swigert 1978, 1988). Because these stereotypes are widespread in society, they are likely to be used not only by members of the juvenile justice system, but also by parents, teachers, and peers. While actual deviant behavior increases the likelihood of being labeled a deviant, delinquency is not a necessary condition for being labeled. The "falsely accused" are persons who refrain from deviance but get labeled anyway (Becker 1963). Moreover, the powerless, having fewer cultural and material resources at their disposal, may be more likely to accept deviant labels. Again, the result is a self-fulfilling prophecy: members of disadvantaged groups are labeled delinquent, which alters their self-conceptions and causes them to deviate, thus fulfilling the prophecy of their initial label. Finally, labels are not restricted to deviance. One can be labeled a conformist or a success at conventional activity, which should increase the likelihood of conventional behavior, while decreasing the likelihood of deviance.

Empirical research on labeling theory has produced equivocal results.

While some research has found official labeling to have trivial effects on self-image (Gibbs 1974), especially when prior self-reported delinquency is controlled (Hepburn 1977), other research has found official labels to have effects for some youth (whites and nonserious delinquents) but not others (Ageton and Elliott 1974). In summarizing this research, Jensen (1980) concluded that official labeling may have a greater impact on delinquent self-images and attitudes among those less heavily involved in delinquency. Research on the effect of official labeling on subsequent delinquent behavior has found positive effects on delinquency (Meade 1974), but when prior levels of self-reported delinquency are controlled, the results have been inconsistent (Thomas and Bishop 1984; Ray and Downs 1986). Recently, Hagan and Palloni (1990) found evidence that official labeling of parents and sons interacts to produce greater self-reported delinquency. They conclude that labeling leads to an intergenerational reproduction of a criminal class, which supports the ideas of Mead, Tannenbaum, and Lemert. While this research literature has led some researchers to dismiss labeling theory (Hirschi 1980), others have concluded that attention should focus more on the consequences of informal rather than official labels (Paternoster and Iovanni 1989). Menard and Morse (1984) found support for the latter proposition: perceived informal social labels had substantial effects on delinquency and helped mediate the effect of IQ on delinquency.[8]

Labeling theory can help specify the relationships between background characteristics, the informal labeling process, and delinquency. First, youths who have engaged in delinquent behavior should be more likely to be labeled delinquent by their parents. Second, insofar as parents act on conventional stereotypes of deviance, their appraisals of their children as either deviant or conforming may be influenced by structural conditions that reflect disadvantages. Urban, minority, lower-class, older adolescent youths may be more likely to be labeled by their parents as deviant and less likely labeled as conforming, in part because they engage in more objective deviance. Indeed, parents could act on stereotypes to such an extent that those parents of disadvantaged children are more likely to label their children deviant, regardless of their children's behavior. This

[8] Some indirect evidence on deviance amplification comes from research on delinquent careers, which finds little specialization or escalation in seriousness of offenses (e.g., Wolfgang, Figlio, and Sellin 1972; Kempf 1987). However, labeling theory does not necessarily imply specialization of offenses· stigmatizing the delinquent will make it difficult to remain conventional, which increases the likelihood of committing a variety of offenses. Deviance amplification does imply some escalation in seriousness of offenses from trivial acts to minor deviance to more serious offenses. Recent research finds some evidence of escalating seriousness through the early adult years (Wolfgang, Thornberry, and Figlio 1987).

would constitute strong evidence for a labeling perspective since the parents share the disadvantages of their children but nevertheless still act on conventional stereotypes. Third, parental appraisals of youths as deviant or conforming will influence their further delinquency, primarily by influencing youths' reflected appraisals of self as deviant or conforming.[9]

DATA AND METHODS

The analyses that follow will examine the propositions above that are derived from both a symbolic interactionist theory of the self and a labeling theory of delinquent behavior. Such an examination requires a research design with at least three features. To examine the labeling hypothesis that parental appraisals vary by social-structural variables, a random sample of a heterogenous population is required. To examine the joint relationships between parental appraisals, reflected appraisals, and delinquency, survey data measuring *perceptual* or subjective social psychological concepts is needed. To examine simultaneously the reciprocal effects of delinquency on parental and reflected appraisals, a longitudinal design is necessary.

Data that meet these requisites were collected by Delbert S. Elliott and his colleagues as part of the National Youth Survey (NYS), a longitudinal study of delinquency and drug use (Elliott, Huizinga, and Ageton 1985; Elliott, Huizinga, and Menard 1989). Employing a multistage cluster sampling frame, the NYS obtained a national probability sample of households in the United States in 1976. After several stages of sampling geographic units, 7,998 households were randomly selected, and all 2,360 eligible youths living in the households were included. Seventy-three percent of those youths (1,725) agreed to participate, signed consent forms, and along with one of their parents, completed first-wave interviews in 1977. As a result, the participating youths are reasonably representative of 11–17-year-olds in the United States. My analyses focus on the first three waves of data for male respondents.[10] Attrition over these waves

[9] The present analysis is concerned with informal labels made by an adolescent's parents and whether those appraisals affect delinquency by affecting the adolescent's reflected appraisals. I am unable to examine the proposition of labeling theory that concerns the effects of formal labels made by officials of the juvenile justice system.

[10] This analysis follows much delinquency research and focuses on the 918 males in the sample. If males and females are pooled, the overall pattern of results is unchanged, thus, our results are not biased because of censoring. Nevertheless, some preliminary analyses suggest that separating males and females reveals some interesting interaction effects. A complete analysis and interpretation of such interactions is beyond the scope of the present article.

was remarkably low: 4% in 1978 and 6% in 1979. Elliott, Knowles, and Canter (1981) examined nonparticipation and attrition and concluded that neither compromised the representativeness of the sample.

The NYS used personal interviews to collect self-reports of delinquent behavior, parents' reported appraisals of their child, and youths' reflected appraisals of themselves from the standpoint of parents, friends, and teachers. The content of the appraisals cluster around four substantive dimensions: (1) sociable, measured by "well-liked" and "gets along well with others"; (2) likely to succeed, measured by a single indicator; (3) distressed, measured by "often upset" and "has a lot of personal problems"; and (4) rule violator, measured by "gets into trouble" and "breaks rules."[11] The delinquency inventory was designed to represent the entire range of delinquent acts for which juveniles could be arrested, and includes all but one part 1 offenses of the Uniform Crime Report (UCR), over half of part 2 offenses, and a range of UCR "other offenses." Following Elliott et al. (1985), I use the categorical response sets, which have less skewed distributions. The analyses will focus on a 24-item scale of general delinquency, since our interactionist theory does not specify a priori reasons for examining specific offenses, and, empirically, recent research finds little evidence that delinquents specialize in offenses. To check the robustness of the results, however, I will also examine three subscales of delinquency: drug use, minor delinquency, and UCR index offenses (see App. A). Finally, the NYS also includes measures of background characteristics relevant to labeling hypotheses: age, race, urban residence, broken home, and family income. Descriptions of the measures appear in Appendix B.

To analyze these data, I first specify measurement models of the reflected appraisals process, and second, incorporate this model into a structural model of the causes and consequences of reflected appraisals. The measurement models allow me to test specific hypotheses about the structure underlying the indicators of reflected appraisals and to estimate and control for the biasing effects of response error. Both measurement and substantive models are estimated simultaneously using Jøreskog and Sorbom's (1988a) LISREL 7 program. Under the assumption of large samples and multivariate normality, this program provides maximum-likelihood estimates, asymptotic standard errors, and a likelihood-ratio test statistic for covariance structure models.

[11] The indicators of distressed and rule violator are selected from a larger set of measures. I originally estimated a measurement model that included, as measures of distress, "are messed up" and "need help" and as measures of rule violator, "are a bad kid" and "do things that are against the law." For parsimony, I retained those indicators that had the best measurement properties.
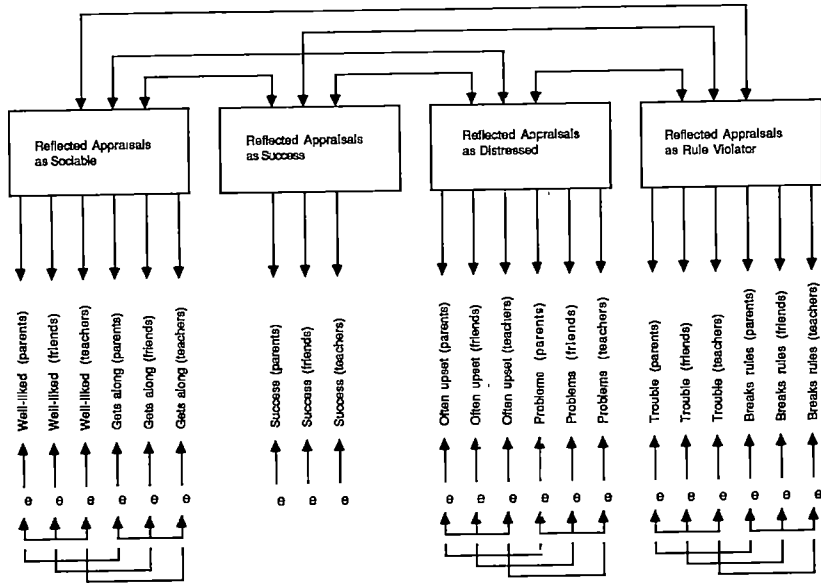
FIG. 2 —A measurement model of reflected appraisals of self

ANALYSIS OF THE MEASUREMENT MODEL

I specify measurement models of youth-reflected appraisals of self as well as parent appraisals of their children. Figure 2 presents a measurement model of reflected appraisals of self made by youths. This model considers each observed indicator as a linear combination of a latent unobserved factor plus a random measurement error. Substantively, the model, based on the theoretical framework above as well as on some preliminary exploratory analyses, implies that youth-reflected appraisals from parents, teachers, and peers coalesce into a single self representing convergence or consensus in reflected appraisals, rather than splitting into conflicting, compartmentalized selves. I did attempt to fit a model that separated reflected appraisals into parent, teacher, and peer factors, but this model was clearly inconsistent with the data. Nevertheless, I did find significant correlations among indicators of a given construct that referred to the same significant other, such as a parent. Adding these 29 error correlations improved the fit from $L^2 = 1248.43$; $df = 463$ to $L^2 = 777.02$; $df = 434$, which is an acceptable fit (Jøreskog and Sorbom's 1988a goodness-of-fit index is .951 and adjusted goodness-of-fit index is .824).[12]

[12] Specifically, I added measurement error correlations in three steps First, I added 18 correlations among identical measures that differed only in who the significant other was (e.g., get in trouble from the standpoint of teachers, parents, and peers). The improvement of fit was $L^2 = 281.20$; $df = 18$ Second, I added nine correlations

Table 1 presents maximum-likelihood estimates of parameters of the measurement model. The indicators of youth-reflected appraisals of self as sociable, as a success, as distressed, and as a rule violator are given in the bottom section of the table. These measures show fairly high validity coefficients, ranging from .47 to .78, particularly given that they are measuring subjective phenomena. In general, with the exception of "personal problems," the parental reflected appraisals are slightly less reliable than those of peers and teachers. While "well-liked" and "gets along well" are equally reliable measures of "sociable," and "gets in trouble" and "breaks rules" are equally reliable measures of "rule violator," "personal problems" is a better measure of "distressed" than is "being upset." Intercorrelations among latent variables range between −.26 and .52, indicating that, as expected, the variables are strongly intercorrelated. Nevertheless, the latent variables are sufficiently distinct to show discriminant validity.

Figure 3 presents a measurement model of parental appraisals of their child. Paralleling the model of youth-reflected appraisals, this model also specifies four latent appraisal factors: parental appraisals of their child as sociable, as a success, as a rule violator, and as distressed. The validity coefficients suggest that the measures are fairly reliable indicators of their theoretical constructs (see table 1). While the measures of rule violator appear equally reliable, "gets along with others" is a more reliable indicator of sociable, while "personal problems" is a more reliable indicator of distress. As with the youth-reflected appraisals, the latent variables underlying parental appraisals are substantially intercorrelated in expected ways, ranging from −.40 to .66.

These results suggest that, for both youth-reflected appraisals and parental appraisals of youths, the indicators are adequate measures of the theoretical constructs, but contain sufficient measurement error—both random and systematic—to require correcting for attenuation due to unreliability.

ANALYSIS OF THE SUBSTANTIVE MODEL

Specifying the Model and Hypotheses

The substantive model, depicted in figure 4, specifies causal relationships among the four following latent constructs: (1) a set of exogenous back-

---

among measures of a given construct that referred to a similar significant other (e.g., rule breaker from the standpoint of teachers and gets into trouble from the standpoint of teachers). The difference in chi-squares was $L^2 = 156.08$; $df = 9$ Finally, I added two additional correlations between getting along with others from the standpoint of teachers, and two other teacher indicators (personal problems and success). The improvement of fit was $L^2 = 34.14$; $df = 2$

TABLE 1

PARAMETER ESTIMATES OF THE MEASUREMENT MODELS. MALES

Variables	Observed Variance	Error Variance	Metric Slope	Validity Coefficient
Parental appraisals.				
SOCIABLE1.				
Well-liked.....................	408	232	.747	657
Gets along	467	.151	1.000[a]	.822
SUCCESS1				
Success	.373	000[a]	1.000[a]	1.000[a]
DISTRESSED1:				
Often upset.....	1 048	.667	1 000[a]	603
Problems	.823	.171	1.307	890
RULE VIOLATOR1.				
Trouble	.744	333	1.000[a]	743
Breaks rules	.989	.496	1.097	.707
Youth-reflected appraisals:				
SOCIABLE2:				
Well-liked (parents)	.354	.276	.946	470
Well-liked (friends)	281	.173	1 112	.619
Well-liked (teachers).............	.334	209	1 200	.613
Gets along (parents)	.333	202	1.227	.628
Gets along (friends)	286	.199	1.000[a]	552
Gets along (teachers)	.299	.216	.957	.516
SUCCESS2·				
Success (parents)	.537	256	910	.724
Success (friends)	565	.224	1.000[a]	.776
Success (teachers)	.514	.228	.917	.746
DISTRESSED2·				
Often upset (parents)........	.982	.666	1.171	.567
Often upset (friends)	741	.508	1.000[a]	557
Often upset (teachers)	.676	406	1.087	635
Problems (parents).....	727	283	1.389	782
Problems (friends).....	.631	305	1.190	.719
Problems (teachers)	542	.210	1.199	.781
RULE VIOLATOR2:				
Trouble (parents)	.812	.400	.955	.714
Trouble (friends)	.788	.334	1.000[a]	.759
Trouble (teachers)	.733	343	.925	.728
Breaks rules (parents)	811	.377	.978	.732
Breaks rules (friends)	.782	.345	.981	748
Breaks rules (teachers)	.706	.294	.951	763

NOTE —N = 851
[a] fixed coefficient
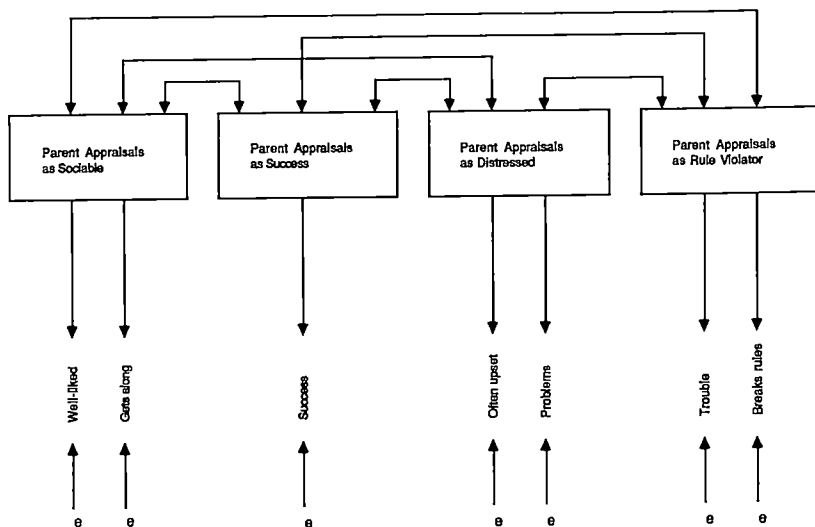
FIG. 3.—A measurement model of parental appraisals of youths
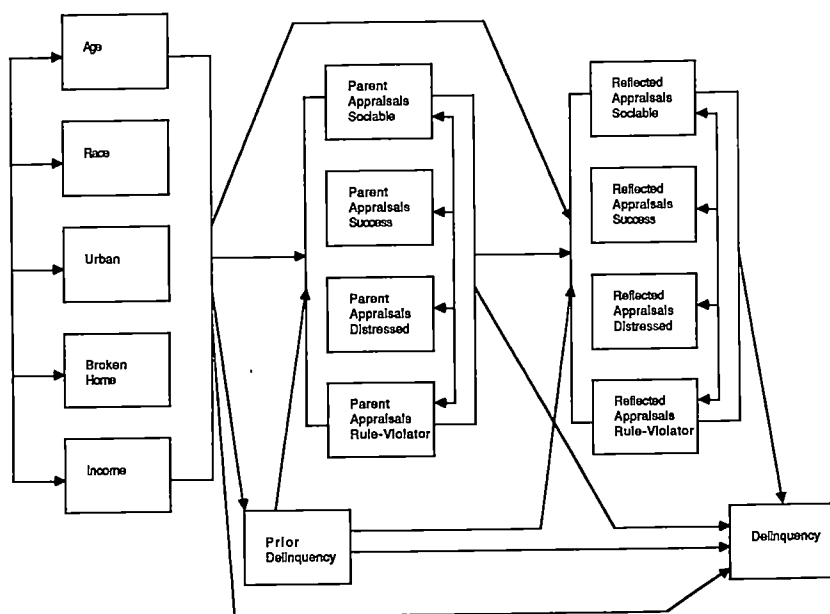
FIG. 4.—A substantive model of parental appraisals, reflected appraisals, and delinquency.

1595

ground variables measured at time 1; (2) a set of endogenous parental appraisals of youths as sociable, as a success, as distressed, and as a rule violator, measured at time 1; (3) a set of endogenous youth-reflected appraisals of self as sociable, as a success, as distressed, and as a rule violator, measured at time 2; and (4) an outcome variable of delinquency, measured at time 3.[13] The causal ordering of the variables follows my theoretical specification: parental appraisals influence reflected appraisals, which in turn influence delinquent behavior. The time ordering of the variables coincides with this causal ordering to reduce ambiguity in making causal inferences. Analyses that vary the precise timing of the variables do not change the substantive findings appreciably (see App. B).

Within this model, we can identify specific hypotheses derived from labeling and symbolic interactionist theories. Consistent with labeling theory, background characteristics reflecting disadvantages should increase the likelihood of negative parental labeling and perhaps decrease the likelihood of positive labeling. Thus, parental labeling of a youth as a rule violator and as distressed may be greater for youths who have committed prior delinquent acts, who are black, and who come from urban, low-income areas and broken homes. As noted above, if disadvantages increase the likelihood of parental labeling, net of prior delinquency, the stereotyping process specified by labeling theorists would receive strong support. Alternatively, it may be that stereotypes are used only by secondary others, such as teachers or juvenile justice officials, and that parents and other significant others use their intimate knowledge of the child in forming appraisals. Finally, parental labeling of the child, particularly as a rule violator, should have substantial total effects on future delinquency. As deviance amplification predicts, youths will commit more crimes if their parents label them as rule violators or as distressed; conversely, they may commit fewer crimes if their parents label them as sociable or as likely to succeed.

Symbolic interactionist theory implies three hypotheses concerning direct effects. First, parental appraisals should have a direct effect on their reflected counterparts, net of prior performance (delinquency). This tests the hypothesis that one's reflected appraisals of self from the standpoint of significant others is, in part, a reflection of the actual appraisals made by those significant others. Second, previous delinquent behavior should exert a direct effect on reflected appraisals of self. Prior delinquency

[13] Relationships among the four parental appraisals are specified as unanalyzed correlations, as are the four youth-reflected appraisals (see fig. 2) These disturbance correlations range from −.36 to .58 for parental appraisals, and from −.15 to .40 for youth-reflected appraisals.

should increase reflected appraisals as a rule violator and as distressed, and perhaps decrease reflected appraisals as sociable and likely to succeed. This hypothesis is consistent with (1) an interactionist perspective that specifies self-images ("me's") as determined in part by prior behavioral resolutions to problematic situations and (2) Bem's (1972) theory of self-perception formation, which posits that individuals form conceptions of self by observing their own behavior. Third, future delinquent behavior should be directly affected by one's reflected appraisals of self. Delinquency should be substantially affected by reflected appraisals of one as a rule violator and perhaps also by reflected appraisals as distressed, sociable, and likely to succeed.

The model also allows us to test several hypotheses that imply parameters constrained to be zero. First is Kinch's (1963) hypothesis that the effect of prior behavior on reflected appraisals is entirely mediated by parental appraisals. Second is Kinch's hypothesis that the effect of prior delinquency on future delinquency is entirely mediated by the intervening reflected appraisals process. In contrast, an interactionist theory would predict that, through habit or nonreflective behavior, prior delinquency will maintain a direct effect on future delinquency. Third is the hypothesis, consistent with Kinch and symbolic interactionism in general, that parental appraisals influence delinquency only indirectly through their effects on reflected appraisals. The alternative hypothesis is that parental appraisals are sufficiently accurate to predict delinquency, even holding prior delinquency and reflected appraisals constant. Fourth is the hypothesis, consistent with labeling and interactionist theories, that the effect of background variables on delinquency works primarily indirectly through the labeling and reflected appraisals process.

Estimation of the Model

Table 2 presents unstandardized parameter estimates of the substantive model; standardized counterparts appear in table 3. The hypotheses derived from labeling theory involve effects of background characteristics and prior delinquent behavior on parental appraisals of youth (tables 2 and 3, rows 2–5). As hypothesized, older youths, urban dwellers, and youths from broken homes, all commit more delinquent acts on average. Consistent with labeling theory, prior delinquent behavior substantially increases parental appraisals of a youth as a rule violator and distressed, while also slightly reducing their appraisals of a youth as sociable or likely to succeed (see rows 2–5 in col. 6). Also consistent with labeling theory, the background variables exert some effect on parental appraisals, particularly rule violator ($R^2 = .13$). Parents of youths who are younger, nonwhite, and from urban areas are more likely to label their

TABLE 2

UNSTANDARDIZED PARAMETER ESTIMATES OF THE SUBSTANTIVE MODEL: MALES

DEPENDENT VARIABLES	PREDETERMINED VARIABLES														R^2
							Parental Appraisals				Youth-reflected Appraisals				
	AGE (1)	RACE (2)	URBAN (3)	BROKEN HOME (4)	INCOME (5)	PRIOR DELINQUENCY1 (6)	SOCIABLE1 (7)	SUCCESS1 (8)	DISTRESS1 (9)	RULE VIOLATOR1 (10)	SOCIABLE2 (11)	SUCCESS2 (12)	DISTRESS2 (13)	RULE VIOLATOR2 (14)	
1 DELINQUENCY1	.022 (.004)	048 (025)	080 (020)	069 (020)	-.004 (004)										066
Parental appraisals:															
2 SOCIABLE1	022 (012)	-018 (067)	016 (054)	-114 (054)	009 (011)	-268 (092)									032
3 SUCCESS1	008 (011)	-040 (063)	001 (051)	-067 (050)	011 (010)	-270 (086)									019
4 DISTRESS1	001 (012)	-020 (069)	-001 (056)	160 (056)	-017 (011)	513 (100)									077
5 RULE VIOLATOR1	-056 (013)	-207 (077)	090 (062)	031 (061)	-004 (012)	708 (107)									126
Youth-reflected appraisals:															
6 SOCIABLE2	000 (007)	-036 (037)	027 (030)	-005 (030)	003 (006)	-004 (053)	097 (039)	018 (026)	005 (036)	-037 (041)					.068
7 SUCCESS2	-017 (012)	166 (067)	059 (054)	049 (053)	032 (011)	-018 (096)	053 (069)	074 (048)	034 (066)	-180 (075)					101
8 DISTRESS2	001 (010)	-138 (054)	-038 (043)	016 (043)	-024 (009)	183 (077)	-072 (055)	-042 (038)	117 (053)	065 (059)					164
9 RULE VIOLATOR2	032 (013)	128 (071)	-056 (056)	016 (056)	-019 (011)	907 (104)	-100 (073)	029 (050)	-120 (070)	379 (083)					285
Delinquent behavior:															
10 DELINQUENCY3 .	018 (004)	040 (025)	029 (020)	007 (020)	-006 (004)	451 (038)	053 (026)	-019 (017)	-010 (025)	084 (029)	166 (046)	006 (019)	-056 (026)	157 (021)	464

NOTE.—SEs are in parentheses, $N = 851$

TABLE 3

STANDARDIZED PARAMETER ESTIMATES OF THE SUBSTANTIVE MODEL: MALES

DEPENDENT VARIABLES	AGE (1)	RACE (2)	URBAN (3)	BROKEN HOME (4)	INCOME (5)	PRIOR DELIN- QUENCY1 (6)	Parental Appraisals				Youth-reflected Appraisals			
							SOCIABLE1 (7)	SUCCESS1 (8)	DISTRESS1 (9)	RULE VIOLATOR1 (10)	SOCIABLE2 (11)	SUCCESS2 (12)	DISTRESS2 (13)	RULE VIOLATOR2 (14)
1 DELINQUENCY1	170	071	139	127	-035									
Parental appraisals														
2 SOCIABLE1	076	-012	012	093	037	-119								
3 SUCCESS1	025	-024	001	-050	042	-110								
4 DISTRESS1	003	-012	-001	118	-061	207								
5 RULE VIOLATOR1	-168	-120	061	022	-013	275								
Youth-reflected appraisals:														
6 SOCIABLE2	-002	-045	040	-008	024	-003	185	037	010	-080				
7 SUCCESS2	-056	105	044	038	125	-008	051	077	036	-197				
8 DISTRESS2	003	-106	-034	015	-112	095	-085	-054	151	086				
9 RULE VIOLATOR2	092	070	-036	011	-065	335	-083	026	-110	360				
Delinquent behavior:														
10 DELINQUENCY3	119	051	042	011	-.044	382	101	-039	-021	184	167	011	-092	.361

NOTE —$N = 851$.

children as rule violators. This is primarily because such youths have committed delinquent acts in the past.

But the only evidence that the disadvantaged may be falsely accused by parents is a countervailing effect of race. Blacks are less likely to be labeled rule violators because they commit fewer delinquent acts; however, net of delinquency, they are more likely to be negatively labeled. Parents in nonintact families are less likely to label their children sociable and more likely to label them distressed (tables 2 and 3, col. 4). This second effect works partly indirectly through prior delinquent behavior. In general, then, these results provide limited support for the labeling hypothesis that, net of their delinquent behavior, youths from disadvantaged backgrounds are more likely to be labeled negatively.

Turning to the reflected-appraisals equations, we find support for the interactionist proposition that reflected appraisals of self are partly a reflection of the objective appraisals made by others. With one exception, parental appraisals of a youth have significant effects on youths' corresponding reflected appraisals of self. This effect is particularly strong for the rule-violator appraisals, perhaps because of the salience of youths' deviant behaviors to parents, who are likely to be concerned about such behavior. The one exception is the coefficient relating parental and reflected appraisals of who is likely to succeed, which does not quite reach significance.[14] The small size of this effect is perhaps due to the nebulous nature of the category, "likely to be a success," especially for adolescents, whose success has yet to be determined.[15]

We also find support for the symbolic interactionist hypothesis that prior behavior influences reflected appraisals indirectly by influencing significant others' appraisals. Prior delinquency has a large total effect (standardized coefficient of .42) on reflected appraisals as rule violator, and a moderate total effect (.17) on distressed. About 25% of the effect of prior delinquency on reflected appraisals as rule violator is mediated by

[14] The nonsignificance of this coefficient could be due to random measurement error in the measure of parental appraisals of a youth as "likely to succeed." Such error, fixed in this model to have a zero variance, could exert a downward bias on this coefficient. A sensitivity analysis on the measurement error variance reveals that increasing the measurement error variance increases the estimated structural coefficient, but also increases its standard error, so that it remains nonsignificant. We would need a larger sample, or perhaps additional indicators, to rule out chance in this estimate. Because the substantive results were not altered when the measurement error variance was fixed to be small, I chose to leave the variance fixed at zero

[15] The bivariate correlations between parental appraisals and youth-reflected appraisals are of course larger, since they do not condition on the background variables, or the effects of parental appraisals of one characteristic on reflected appraisals of another. The bivariate correlations are .244 for sociable, .183 for success, .316 for distress, and .388 for rule violator.

parental appraisals as a rule violator. Nevertheless, contrary to Kinch's mediation hypothesis, even holding parental appraisals constant, prior delinquency significantly influences reflected appraisals as rule violator (standardized coefficient of .34) and as distressed (.09). This finding supports an interactionist theory, as well as Bem's (1972) self-perception theory.

Overall, the model explains nearly half of the variance in delinquent behavior (tables 2 and 3, row 10). Of the youth-reflected appraisal variables, rule violator has by far the largest effect (a standardized coefficient of .36). Thus, persons who perceive that others view them as one who violates rules, or gets in trouble, engage in more delinquent acts. This supports the major hypothesis of an interactionist theory of delinquency: behavior is strongly influenced by reflected appraisals. Of the remaining reflected appraisal variables, sociable has a substantial positive effect on delinquency, and distressed a small negative effect. Net of the other variables in the model, youths who see themselves (from the standpoint of others) as sociable commit more delinquent acts, while those who see themselves as distressed commit slightly fewer delinquent acts.

Consistent with a deviance amplification hypothesis, parental appraisals of a youth as a rule violator have a substantial total effect on delinquency (a standardized coefficient of .29). Moreover, nearly one-half of this effect is mediated by youth-reflected appraisals as a rule violator. This supports an important hypothesis of symbolic interactionism: parental appraisals influence youth-reflected appraisals, which in turn influence delinquency. Nevertheless, even controlling for youth-reflected appraisals, parental appraisals of youth as a rule violator exert a substantial effect on delinquency. It appears that parents' appraisals of youths are, to some extent, more accurate than youths' perceptions of those appraisals. This finding fails to support a symbolic interactionist perspective. Parental appraisals of youths as sociable also exert a significant total effect on delinquency; however, only a small portion is mediated by youth-reflected appraisals. This effect is positive, meaning that, net of the other explanatory variables, more likeable and sociable youths commit more delinquent acts. There is a small but significant indirect effect of parental appraisals as distressed on delinquency through reflected appraisals as distressed.

Prior delinquency (measured at time 1) has a very large total effect on delinquency at time 3 (standardized effect of .55), about 30% of which is mediated by parental and reflected appraisals as a rule violator. This implies that males between the ages of 13 and 19 are fairly stable in delinquent behavior: those who engage in delinquency are likely to have engaged in delinquency two years earlier. This finding also implies that, as predicted by an interactionist theory of the self, part of the stability

in delinquency is due to reflected appraisals as a rule violator and part is due to habit (nonreflective behavior). Alternatively, the direct effect could reflect unmeasured mechanisms not included in the model. Finally, with one exception, the intervening mechanisms specified by the labeling and symbolic interaction process account for most of the effects on delinquency of the background variables. The one exception is age—about one-half of the total effect of age on delinquency is direct.

SUMMARY AND CONCLUSIONS

In sum, these analyses yield six principal findings, which provide general support for a symbolic interactionist conceptualization of reflected appraisals and delinquency. First, youths' reflected appraisals of themselves from the standpoint of parents, teachers, and friends coalesce into a consensual self, rather than remaining compartmentalized as distinct selves. This holds for reflected appraisals as rule violators, distressed, sociable, and likely to succeed. Second, consistent with labeling theory, parental labels of youths as rule violators are more likely among delinquents, nonwhites, and urban dwellers. Most of these effects operate indirectly through prior delinquency; thus, we find only modest evidence of disadvantaged youths being falsely accused. This, however, is not surprising, since the labels investigated are those of parents, who share the disadvantages of the youth. A stronger test of this proposition would examine formal labels of secondary others, such as the juvenile justice system.

Third, with the exception of likely to succeed, youths' reflected appraisals of themselves are strongly influenced by their parents' independent appraisals of them. This result is particularly strong for the rule violator variable, the reflected appraisal most relevant to labeling and delinquency. This finding, at least with regard to rule violator, suggests that youths accurately perceive their parents' appraisals of them and that the reflected-appraisal constructs are capturing meaningful elements of self-concept. Moreover, this supports the proposition of symbolic interactionism that reflected appraisals of the self arise through role-taking in transactions and, therefore, are in part a reflection of others' actual appraisals.

Fourth, previous delinquent behavior influences reflected appraisals of self. Consistent with predictions from labeling theory, this effect works partly indirectly through parental appraisals. Prior delinquency, however, also affects reflected appraisals directly. This implies that reflected appraisals are not a mirror reflection of others' appraisals, as implied by a literal interpretation of Cooley's (1922) looking-glass self, but rather are the result of selective perception of others' appraisals and of previous

behavior. Fifth, as predicted by symbolic interactionism, reflected appraisals as a rule violator exert a large effect on delinquent behavior and mediate much of the effect of parental appraisals as a rule violator on delinquency. Contrary to interactionism, however, parental appraisals as a rule violator still exert a direct effect on delinquency. Sixth, age, race, and urban residence exert significant total effects on delinquency, most of which work indirectly through prior delinquency, and partially through the rule-violator reflected appraisal.

These findings suggest that an interactionist conception of the self as reflected appraisals provides an important cause and consequence of delinquent behavior. The results, however, find Kinch's (1963) theoretical model of reflected appraisals and behavior to be overly restrictive. In particular, prior delinquent behavior influences reflected appraisals even when parental appraisals are held constant; prior delinquent behavior affects delinquency even holding both parental and reflected appraisals constant; and parental appraisals as a rule violator influence delinquency even holding reflected appraisals as a rule violator constant. One could argue that these results contradict only a literal interpretation of Kinch's models, and, if viewed loosely, the results are generally consistent with his arguments. One could also argue that the model estimated here does not include all possibly relevant appraisals by others and reflected appraisals, and if it did, these results might support Kinch's model. Moreover, the present model could not consider Kinch's hypothesis that the effects of reflected appraisals on delinquency are not direct, but are instead mediated by self-appraisals.

Viewed more broadly, these results indicate that the concept of role-taking, as specified by symbolic interactionism, is important for delinquency. Unlike most previous research on the self and delinquency, which examines global self-esteem, I have examined specific meanings of the self pertaining to violating rules. I found that one aspect of the self—reflected appraisals as a rule violator—has strong effects on delinquency. Furthermore, an interactionist perspective would imply that other specific meanings of the self may also influence delinquent behavior. These include specific attitudes and motives concerning delinquent and conforming alternatives, the specific reactions to and attitudes about delinquency held by significant others, and the presentation of situational motives for delinquency by significant others.

Such mechanisms could account for the finding that parental appraisals of a youth as a rule violator affect delinquency, net of the youth's reflected appraisals. While the finding could mean that parents are particularly accurate in predicting the behavior of their children, as noted above, it could also mean that parental appraisals work indirectly through an aspect of the self besides reflected appraisals, such as self-esteem or antici-

pated reactions. Alternatively, the result may be consistent with a structural explanation. Parents may form their appraisals based in part on the role occupied by the child (as well as the role occupied by the parent). They may consequently alter their treatment of the youth—which affects delinquent behavior—but never communicate their appraisal to the child. For example, youths occupying the role of troublemaker may be viewed as such by parents, which causes parents to be alienated from their child, ultimately leading to increased delinquency. Or a parent with a criminal history may be more likely to view the child as a delinquent, which reduces the affection given the child and thereby increases the likelihood of delinquency. In either case, parental appraisals of the youth may lead to delinquency without their being consciously perceived by the youth. Such mechanisms, if found to exist, would underscore the importance of a structural version of symbolic interaction (Stryker 1980).

Symbolic interactionism can also provide a framework for integrating the literatures on reflected appraisals, global self-esteem, and delinquent behavior. From this perspective, delinquent behavior is determined primarily by specific meanings concerning delinquency, including reflected appraisals, role identities, specific motives concerning delinquency, reactions of significant others to delinquency, and presentation of situational motives by significant others. Accordingly, low global self-esteem and high self-rejection are products of previous social transactions and can motivate the individual to act in ways that increase his or her self-regard (Hewitt 1970; Kaplan 1980). The extent to which this is accomplished by delinquent behavior depends on the outcome of reflective behavior in problematic situations, which in turn is shaped in part by one's stable self-image concerning delinquency. This implies that effects of global self-esteem on delinquency will be dwarfed by effects of specific meanings of the self concerning delinquency and may be conditioned by such specific meanings. Future research is needed to estimate relative weights of such processes and to determine whether conditional effects are involved. Such research would link traditional studies of self-esteem to our findings supporting a symbolic interactionist theory of delinquency.

APPENDIX A

Examination of the Robustness of the Findings

While the findings above provide general support for an interactionist theory of delinquency, they could be artifacts of certain methodological assumptions made in the analysis. Therefore, I performed additional analyses to examine the robustness of the findings to four competing specifications and interpretations. First, in the models reported above,

prior delinquency is lagged two years from the outcome variable of delinquent behavior. This was done to make the causal ordering among prior delinquency, parental appraisals, reflected appraisals, and future delinquency coincide with their temporal ordering of measurement. I did not include delinquency at time 2, which would be causally ambiguous with respect to parental and youth-reflected appraisals. If time-2 delinquency is included, then the effects of reflected appraisals on time-3 delinquency are somewhat attenuated; nevertheless, the effect of rule violator is still substantial in size and statistically significant. However, if parental and reflected appraisals, as well as prior delinquency, are measured at time 1, and future delinquency measured at time 2, then only time-1 delinquency and parental label as a rule violator significantly influence time-2 delinquency. This appears to be due to the high stability in delinquency between time 1 and time 2.

Second, the results could fail to apply to measures of delinquency other than the 24-item index used. Therefore, I estimated the model substituting three conceptually relevant indexes of delinquency: drug offenses, minor delinquency, and index offenses. Overall, with some minor exceptions, these models yield the same substantive story.

Third, a potential threat to the validity of causal inferences concerns the relationship between reflected appraisals as a rule violator and delinquency. Even though I have controlled for prior levels of delinquency, it still could be that rule violator is serving as a proxy for minor forms of prior deviance, rather than a reflected appraisal. In other words, the effect of rule violator on delinquency could be spurious owing to the omission of prior minor deviant behavior. Fortunately, some additional minor deviant acts were measured at time 1 (but not time 2), including vandalized property of family, school, and other; threw objects at people; lied about age to gain entrance; sold marijuana; cheated on school tests; illegally hitchhiked; stole from family; bought liquor for a minor; avoided paying for things; was publicly drunk; stole at school; and skipped classes. I first added these 10 items to the index of prior delinquency, which yielded a 34-item index, reestimated the model, and obtained identical results. I then estimated a model that included both measures of prior delinquency (the 24-item index plus the 10-item scale) and again obtained identical results. Thus, it seems reasonable to conclude that reflected appraisals as a rule violator is not serving as a proxy for prior minor delinquency—at least not as measured here.

Finally, it could be that the results are methodological artifacts of two assumptions: the dependent variables are measured on unbounded interval scales and the observed variables are distributed multinormally. Although Monte Carlo evidence suggests the LISREL estimator is robust to these assumptions, I nevertheless estimated a model that uses the

log of delinquency, which reduces skewness and kurtosis, and obtained identical results. I then used Muthen's (1984) general framework to estimate a model that assumed that delinquency is a continuous variable, but all other measures are ordinal. Specifically, I used Jøreskog and Sorbom's (1988a, 1988b) approach, which uses PRELIS to estimate polychoric and polyserial correlations and their asymptotic covariances, then uses weighted least squares to obtain asymptotic distribution-free estimates of parameters and standard errors. Given the sample size, however, I could only estimate the important core components of the model: the relationships between background variables, parent appraisals as a rule violator, reflected appraisals as a rule violator, and delinquency. This analysis gives similar overall results, although effects are somewhat attenuated: reflected appraisals still have significant effects on delinquency, but parental appraisals do not.

APPENDIX B

Descriptions of Measures

Background Characteristics and Delinquent Behavior

AGE	Years of age of the youth respondent.
RACE	Race of the youth respondent (0 = black; 1 = non-black).
URBAN	Urbanicity (0 = rural or suburban; 1 = urban).
INCOME	Family Income (10-point scale in $4,000 increments: 1 = $6,000 or less; 10 = more than $38,000).
BROKEN HOME	Broken home (0 = intact; 1 = at least one parent absent).
DELINQUENCY1	Index of 24 delinquent acts committed in the past year, including the following: auto theft, $5 theft, $5–$50 theft, $50 theft, buying stolen goods, runaway, concealed weapon, aggravated assult, prostitution, sexual intercourse, gang fights, sold marijuana, hit parents, hit teacher, hit students, disorderly conduct, sold drugs, joyriding, sexual assault, strong-armed students, strong-armed teachers, strong-armed others, breaking and entering, panhandled.
DELINQUENCY3	Index of 24 delinquent acts committed between years 2 and 3.

1606

Parental Appraisals

For the following measures, the parent was asked, "Please listen carefully and tell me how much you agree or disagree with each of the words or phrases as a description of your son or daughter."

SOCIABLE1

| Well liked | "My son or daughter is well liked." |
| Gets along | "My son or daughter gets along well with others." |

DISTRESSED1

| Often upset | "My son or daughter is often upset." |
| Problems | "My son or daughter has a lot of personal problems." |

SUCCESS1

| Success | "My son or daughter is likely to succeed." |

RULE VIOLATOR1

| Trouble | "My son or daughter gets into trouble." |
| Breaks rules | "My son or daughter breaks rules." |

Youth-Reflected Appraisals

For the following measures, the youth was asked, "I'd like to know how your parents, friends, and teachers would describe you. I'll read a list of words or phrases and for each, will ask you to tell me how much you think your parents would agree with that description of you. I'll repeat the list twice more, to learn how your friends and your teachers would describe you."

SOCIABLE2

| Well liked | "Parents agree I am well liked."
"Friends agree I am well liked."
"Teachers agree I am well liked." |
| Gets along | "Parents agree I get along well with others."
"Friends agree I get along well with others."
"Teachers agree I get along well with others." |

DISTRESSED2

| Often upset | "Parents agree I am often upset."
"Friends agree I am often upset."
"Teachers agree I am often upset." |
| Problems | "Parents agree I have a lot of personal problems." |

"Friends agree I have a lot of personal problems."
"Teachers agree I have a lot of personal problems."

SUCCESS2

Success
"Parents agree I am likely to succeed."
"Friends agree I am likely to succeed."
"Teachers agree I am likely to succeed."

RULE VIOLATOR2

Trouble
"Parents agree I get into trouble."
"Friends agree I get into trouble."
"Teachers agree I get into trouble."

Breaks rules
"Parents agree I break rules."
"Friends agree I break rules."
"Teachers agree I break rules."

REFERENCES

Ageton, Suzanne, and Delbert S. Elliott. 1974. "The Effects of Legal Processing on Self-Concepts." *Social Problems* 22.87–100.

Bachman, Jerald G., Patrick M O'Malley, and Jerome Johnston. 1978. *Youth in Transition*. Vol. 6, *Adolescence to Adulthood: Change and Stability in the Lives of Young Men* Ann Arbor, Mich.: Institute for Social Research.

Becker, Howard S. 1963. *Outsiders: Studies in the Sociology of Deviance*. New York: Free Press.

Bem, Daryl J. 1972. "Self-Perception Theory." Pp. 2–62 in *Advances in Experimental Social Psychology*, vol. 6. Edited by Leonard Berkowitz. New York: Academic Press.

Blumer, Herbert. 1969. *Symbolic Interactionism: Perspective and Method*. Englewood Cliffs, N.J.. Prentice-Hall

Briar, Scott, and Irving Piliavin. 1965. "Delinquency, Situational Inducements, and Commitment to Conformity." *Social Problems* 13:35–45.

Bynner, John M., Patrick M. O'Malley, and Jerald G. Bachman. 1981. "Self-Esteem and Delinquency Revisited." *Journal of Youth and Adolescence* 10:407–44.

Cohen, Albert. 1955. *Delinquent Boys*. Glencoe, Ill.: Free Press.

Cooley, Charles H. 1922. *Human Nature and the Social Order,* rev. ed. New York: Scribners.

Dewey, John. 1922 *Human Nature and Conduct* New York: Modern Library.

Elliott, Delbert S., Suzanne Ageton, and Rachelle J. Canter. 1979. "An Integrated Theoretical Perspective on Delinquent Behavior." *Journal of Research in Crime and Delinquency* 16:3–27.

Elliott, Delbert S., David Huizinga, and Suzanne Ageton. 1985. *Delinquency and Drug Use*. Beverly Hills, Calif.: Sage.

Elliott, Delbert S., David Huizinga, and Scott Menard. 1989 *Multiple Problem Youth: Delinquency, Substance Use, and Mental Health Problems*. New York: Springer-Verlag.

Elliott, Delbert S , Brian A. Knowles, and Rachelle J Canter. 1981. *The Epidemiology of Delinquent Behavior and Drug Use among American Adolescents*. Project

Report no. 14, A National Youth Survey. Boulder, Colo Behavioral Research Institute.

Farrell, Ronald A., and Victoria Lynn Swigert. 1978. "Prior Offense Record as a Self-Fulfilling Prophecy." *Law and Society Review* 12:437–53.

———— 1988. "A Hierarchical Systems Theory of Deviance." Pp. 391–407 in *Social Deviance*, 3d ed. Edited by Ronald A. Farrell and Victoria L. Swigert. Belmont, Calif.: Wadsworth.

Felson, Richard B. 1980 "Communication Barriers and the Reflected Appraisal Process " *Social Psychology Quarterly* 43·223–32.

———— 1985. "Reflected Appraisal and the Development of Self." *Social Psychology Quarterly* 48.71–77.

————. 1989. "Parents and the Reflected Appraisal Process: A Longitudinal Analysis." *Journal of Personality and Social Psychology* 56.965–71.

Festinger, Leon. 1954. "A Theory of Social Comparison Processes." *Human Relations* 7.117–40.

Gibbons, Donald C. 1971. "Observations on the Study of Crime Causation." *American Journal of Sociology* 77:262–78.

Gibbs, Leonard E 1974. "Effects of Juvenile Legal Procedures on Juvenile Offenders' Self-Attitudes." *Journal of Research in Crime and Delinquency* 11:51–55.

Glaser, Daniel. 1979 "A Review of Crime-Causation Theory and Its Application." Pp. 203–37 in *Crime and Justice: An Annual Review of Research*, edited by Norval Morris and Michael Tonry. Chicago: University of Chicago Press.

Hagan, John, and Alberto Palloni. 1990. "The Social Reproduction of a Criminal Class in Working-Class London, circa 1950–1980." *American Journal of Sociology* 96:265–300.

Hepburn, John R. 1977. "The Impact of Police Intervention upon Juvenile Delinquents." *Criminology* 15:235–62.

Hewitt, John P. 1970. *Social Stratification and Deviant Behavior*. New York: Random House.

————. 1988 *Self and Society*, 4th ed Boston: Allyn & Bacon.

Hirschi, Travis. 1980. "Labeling Theory and Juvenile Delinquency: An Assessment of the Evidence." Pp. 181–204 in *The Labelling of Deviance: Evaluating a Perspective*, edited by Walter R. Gove. New York: Wiley.

Hyman, Herbert H., and Eleanor Singer. 1968. *Reading in Reference Group Theory and Research*. New York: Free Press

Jensen, Gary F. 1980. "Labeling and Identity· Toward a Reconciliation of Divergent Findings." *Criminology* 18·121–29

Jøreskog, Karl G., and Dag Sorbom. 1988*a*. *LISREL 7: A Guide to the Program and Applications*. Chicago: SPSS, Inc.

————. 1988*b*. *PRELIS: A Program for Multivariate Data Screening and Data Summarization A Preprocessor for LISREL*, 2d ed. Mooresville, Ind.· Scientific Software, Inc.

Kaplan, Howard B. 1975. *Self-Attitudes and Deviant Behavior*. Pacific Palisades, Calif.: Goodyear.

————. 1980. *Deviant Behavior in Defense of Self*. New York: Academic Press.

Kaplan, Howard B., Robert J. Johnson, and Carol A. Bailey. 1987. "Deviant Peers and Deviant Behavior. Further Elaboration of a Model." *Social Psychology Quarterly* 50:277–84.

Kempf, Kimberly L. 1987. "Specialization and the Criminal Career." *Criminology* 25:399–420.

Kinch, J. W. 1963. "A Formalized Theory of the Self-Concept." *American Journal of Sociology* 68:481–86.

Kitsuse, John I., and Aaron V Cicourel. 1963. "A Note on the Official Uses of Statistics." *Social Problems* 2.131–39

Lemert, Edwin. 1951 *Social Pathology* New York. McGraw-Hill
———. 1972. *Human Deviance, Social Problems, and Social Control*, 2d ed. Englewood Cliffs, N.J : Prentice-Hall.
McCall, George J., and J L Simmons. 1978. *Identities and Interaction*. New York: Free Press.
McCarthy, John D , and Dean R. Hoge. 1984 "The Dynamics of Self-Esteem and Delinquency." *American Journal of Sociology* 90:396–410.
Mead, George H. 1934. *Mind, Self and Society*. Chicago: University of Chicago Press.
———. 1938. *The Philosophy of the Act*. Chicago. University of Chicago Press
———. (1918) 1964. "The Psychology of Punitive Justice." Pp. 212–39 in *Selected Writings: George Herbert Mead*, edited by Andrew J. Reck. Chicago: University of Chicago Press.
Meade, Anthony C. 1974 "The Labeling Approach to Delinquency State of the Theory as a Function of Method." *Social Forces* 53·83–91.
Menard, Scott, and Barbara J. Morse 1984. "A Structuralist Critique of the I Q.-Delinquency Hypothesis: Theory and Evidence." *American Journal of Sociology* 89.1347–78
Miyamoto, S. Frank, and Sanford Dornbusch. 1956. "A Test of the Symbolic Interactionist Hypothesis of Self-Conception." *American Journal of Sociology* 61:399–403
Muthen, Bengt 1984. "A General Structural Equation Model with Dichotomous, Ordered, Categorical, and Continuous Latent Variable Indicators." *Psychometrika* 49:115–32.
Paternoster, Raymond, and LeeAnn Iovanni. 1989. "The Labeling Perspective and Delinquency: An Elaboration of the Theory and Assessment of the Evidence " *Justice Quarterly* 6:359–94.
Ray, Melvin C., and William R. Downs. 1986. "An Empirical Test of Labeling Theory Using Longitudinal Data " *Journal of Research in Crime and Delinquency* 23:169–94.
Reckless, Walter C., Simon Dinitz, and Ellen Murray. 1956 "Self Concept as an Insulator against Delinquency." *American Sociological Review* 21:744–46.
Rosenberg, Florence R., and Morris Rosenberg. 1978 "Self-Esteem and Delinquency." *Journal of Youth and Adolescence* 7.279–91.
Rosenberg, Morris. 1979 *Conceiving the Self* New York. Basic Books.
Rosenberg, Morris, Carmi Schooler, and Carrie Schoenbach. 1989. "Self-Esteem and Adolescent Problems. Modeling Reciprocal Effects." *American Sociological Review* 54.1004–18.
Rosenberg, Morris, and Roberta G Simmons. 1972. *Black and White Self-Esteem: The Urban School Child*. Washington, D.C.: American Sociological Association
Schrauger, J Sydney, and Thomas J. Schoeneman. 1979. "Symbolic Interactionist View of Self· Through the Looking Glass Darkly." *Psychological Bulletin* 86:549–73.
Schwartz, Michael, and Sheldon Stryker 1970. *Deviance, Selves, and Others*. Washington, D C. American Sociological Association.
Shibutani, Tamotsu. 1961. *Society and Personality*. Englewood Cliffs, N.J.: Prentice-Hall.
Short, James F., Jr. 1974. "Collective Behavior, Crime and Delinquency " Pp. 403–49 in *Handbook of Criminology*, edited by D. Glaser Chicago Rand McNally.
Short, James F., Jr., and Fred L. Strodtbeck 1965. *Group Process and Gang Delinquency*. Chicago: University of Chicago Press.
Simmons, J L 1965 "Public Stereotypes of Deviants." *Social Problems* 13:223–32
Smelser, Neil J 1963. *Theory of Collective Behavior*. New York. Free Press.
Stryker, Sheldon. 1980. *Symbolic Interactionism*. Menlo Park, Calif.: Benjamin/Cummings

Stryker, Sheldon, and Elizabeth A. Craft. 1982. "Deviance, Selves, and Others Revisited." *Youth and Society* 14:159–83.

Tannenbaum, Frank. 1938. *Crime and the Community*. Boston: Ginn

Thomas, Charles W., and Donna M. Bishop. 1984 "The Effect of Formal and Informal Sanctions on Delinquency: A Longitudinal Comparison of Labeling and Deterrence Theories." *Criminology* 75 1222–45.

Turner, Ralph H. 1964. "Collective Behavior." Pp. 382–425 in *Handbook of Modern Sociology,* edited by Robert E L. Faris. Chicago: Rand McNally

Wells, L. Edward. 1978. "Theories of Deviance and the Self-Concept." *Social Psychology Quarterly* 41:189–204.

Wells, L. Edward, and Joseph H. Rankin. 1983. "Self-Concept as a Mediating Factor in Delinquency." *Social Psychology Quarterly* 46:11–22.

Wolfgang, Marvin G., Robert Figlio, and Thorsten Sellin. 1972. *Delinquency in a Birth Cohort* Chicago: University of Chicago Press.

Wolfgang, Marvin G., Terrance P. Thornberry, and Robert M. Figlio. 1987. *From Boy to Man, From Delinquency to Crime*. Chicago. University of Chicago Press.

Wrong, Dennis H. 1961. "The Oversocialized Conception of Man in Modern Sociology." *American Sociological Review* 26:183–93.

Successful Aging: A Life-Course Perspective on Women's Multiple Roles and Health[1]

Phyllis Moen, Donna Dempster-McClain, and Robin M. Williams, Jr.
Cornell University

Panel data from a sample of 313 women who were wives and mothers in 1956 and were interviewed both in 1956 and in 1986 are used to consider the pathways that lead to health and social integration. Possible relationships were explored between the number, duration, timing, and episodes of various nonfamily roles throughout adulthood and subsequent health and multiple-role occupancy. It was found that occupying multiple roles in 1956, participating in volunteer work on an intermittent basis, and belonging to a club or organization were positively related to various measures of health and that occupying multiple roles in 1956, as well as doing volunteer work, was positively related to occupying multiple roles in 1986.

The process of aging is frequently depicted as synonymous with the onset of poor health and disability. Still, as Riley and Bond (1983) point out, "human beings do not age as specimens in laboratories," but, rather, in complex ways that include social and psychological processes as much as biological ones. Reduced activity levels and sickness may often be concomitant with aging, but the relationship between growing older and poor health is not immutable. Considerable numbers of individuals age successfully with few health problems or disabilities, while others face

such difficulties quite early in life. Most important, many people deal well with particular afflictions, while others with the same difficulties become functionally impaired. While there are diverse definitions of health, we prefer to consider health as a multidimensional continuum, rather than a dichotomy, and in terms of functional abilities and the absence of chronic disease in addition to subjective assessments of one's fitness (Antonovsky 1987).

One promising corollary, and possible precursor, of health in later adulthood is social integration in the form of involvement in multiple roles. Occupying multiple roles—such as worker, club member, and churchgoer—has been positively linked to health and to longevity (Berkman and Breslow 1983; House, Landis, and Umberson 1988; Moen, Dempster-McClain, and Williams 1989). Such multiple-role occupancy may be especially important in later adulthood, a time when role *reduction,* rather than role *accumulation,* becomes increasingly common in our culture (Morgan 1988).

However, cross-sectional research linking health with social integration, defined as "the existence or quantity of social ties or relationships," is problematic because the direction of effects is unclear (House, Umberson, and Landis 1988, p. 302). As Verbrugge (1983) points out, the issue is one of social causation versus social selection. Social *causation* assumes that social integration (occupying multiple roles) influences health.[2] By contrast, social *selection* assumes that healthy people are the ones most likely to take on and maintain multiple social roles. But the social causation versus social selection argument is less crucial than an understanding of the pathways to health and social integration in later adulthood. Causation and selection are probably both operating simultaneously and interactively in a dynamic cascade of events over the life course. Successful aging probably encompasses *both* social integration (multiple roles) *and* health in the later years of life. What is required is a dynamic approach to health and social integration, to examine the extent to which experiences throughout the life course shape physical abilities and involvement in multiple roles later in life (e.g., Riley and Riley 1989; Rowe and Kahn 1987).

Such an approach is especially important when looking at women's lives, since women's social integration into the larger society has been circumscribed by the primacy of their family obligations (Hughes and Gove 1981). Women are also more likely than men to spend their later adulthood alone, without the presence (and support) of a spouse. Aging women are particularly susceptible to social isolation as they leave or lose

---

[2] In this article the terms multiple roles, multiple-role occupancy, multiple-role involvements, and social integration are used synonymously

employment and other roles in addition to their frequent loss of the wife role through divorce or widowhood (Smith and Moen 1988).[3]

This paper employs panel data to examine various dimensions of role involvement in the prime of life that might be conducive to social integration as well as health in later adulthood. We draw on a sample of women aged 55–81 in 1986 who had been interviewed both in 1956 and, 30 years later, in 1986.

THEORETICAL BACKGROUND

A life-course perspective provides a useful lens with which to view successful aging. From this perspective, the primary issue is not whether multiple-role involvement affects health or vice versa, but what the pathways are to social integration and health in the later years of adulthood. Three life-course themes are relevant: timing, process, and context (see Elder 1985).

Timing relates to the incidence, duration, and sequence of roles throughout the life course. For example, the employment role has been positively related to women's health (Verbrugge 1983; Waldron and Jacobs 1989). But knowing whether or not a woman is employed at any one point in time may be less useful than knowing the duration and patterning of her labor-force participation throughout adulthood.

Process focuses on aging as a series of role transitions rather than as a single event. Some roles are left behind, as in retirement from employment, while others, such as "retiree," are entered. Process, like timing, also draws attention to role trajectories, the way roles are played out over the life course.

Context, in terms of the personal circumstances of women's lives— their education, marital status, and family size, as well as their age—may have important repercussions for their social integration and health in the later years of adulthood. Age, in particular, may reflect not only physiological differences but also unique historical circumstances. Women from different cohorts were born into vastly different worlds, with different expectations regarding women's roles and lives.

Taking into account timing, process, and context, the study of aging and health from a life-course perspective becomes an investigation of pathways, of the connections between different phases of life, and how circumstances in early adulthood may affect health and social integration later in life.

---

[3] However, a reduction in social roles may be more a reflection of increasing selectivity than of disengagement (Fredrickson and Carstensen 1990)

Multiple Roles

Since Durkheim ([1897] 1951), social integration in the form of multiple-role involvements has been promoted as having salutary consequences for health and well-being. But social integration at one life phase may or may not be conducive to social integration and health in the later years of adulthood. For example, family responsibilities may constrain young women's involvement in nonfamily roles, such as that of paid or unpaid worker, for a number of years. Thus, the level of role involvement during the child-rearing phase may have little relationship to women's subsequent social integration or health. On the other hand, social integration during early adulthood, despite the demands of the roles of wife and mother, may signal an active involvement in society that persists into old age.

Previous findings (Berkman and Breslow 1983; House, Robbins, and Metzner 1982; Moen et al. 1989) suggest that multiple roles do have positive health outcomes. This supports a role-enhancement theoretical orientation, suggesting the advantages of role accumulation and multiple identities, as opposed to social isolation (Sieber 1974; Thoits 1983, 1986). But is it multiple-role occupancy at any one point in time that is most predictive of health, or is it the duration, timing, or number of spells (episodes) of particular roles that matters?

Family Roles

Marriage has been positively linked to health, whereas divorce and widowhood have negative health effects (e.g., Verbrugge 1989). A life-course focus draws attention to the *timing* of a marital dissolution. It is not clear, from the existing literature, whether the more critical factor in relation to health is the state of nonmarriage or the recency of the transition from marriage to nonmarriage. Moreover, since widowhood becomes increasingly common in the later years of women's lives, being nonmarried may have fewer negative health effects for older women than for women whose marriages dissolve in earlier adulthood, or for men.

Motherhood is also a common role for women and one from which, unlike marriage, women rarely exit. Little is known about the effect of motherhood on social integration and health in later life. Still, the number of children raised or the duration of adulthood spent in caring for young children could have long-term health consequences. Children in the home may promote health-enhancing regimens as women attempt to establish good examples for their children (Umberson 1987), or they may result in women spending less time and effort on self-care (Gove and Hughes

1979). The long-term repercussions of children for health in the postparental phase has not been adequately investigated; however, Moen et al. (1989) found that the number of children was positively related to longevity among women who were wives and mothers in the 1950s, and Umberson (1987) found that the presence of children in the home deterred negative health behaviors (see also Kobrin and Hendershot 1977). Litwak (1985) suggests that family relationships may be especially important for health later in life.

But family roles carry obligations as well as benefits. One potentially costly role, in terms of women's health and social integration, is that of caregiver to adult relatives. While the care involved in the raising of young children is anticipated, frequently welcomed, and circumscribed in the life span, caring for an infirm parent or a handicapped adult child can become an unexpected but absorptive role that may restrict women's participation in other activities and produce debilitating strains in their lives.

Subjective Well-Being

It may not be the number, timing, or duration of roles but women's *satisfaction* with their role involvements, themselves, and their lives that most influences their health. Women who experience role strain in combining employment or community activities with homemaking and child care may suffer, rather than benefit, from occupying multiple roles. In contrast, women who are satisfied with their lives and who have high levels of self-esteem are more likely than those without these orientations to adopt a healthy life-style. Thus, psychological well-being early in life should have long-term consequences for health in the later years (Aneshensel, Frerichs, and Huba 1984).

Particular Roles

Two nonfamily roles, as an employee in paid work or a volunteer in unpaid work, provide mechanisms for women's active participation in the larger society. Recent research has examined the positive effects of employment for women's health (Verbrugge 1983), but no one has examined the long-term significance of the duration, number of spells, or timing of employment throughout adulthood for health in later life.

Even less is known about the health consequences of volunteer work. In an earlier study Moen et al. (1989) found that role membership in clubs or organizations was positively related to women's longevity. They posited that it may be the degree of choice involved in social participation in community activities that was beneficial. Other factors may be in-

volved as well. Social participation may provide gratifying social involvement and recognition, reduction in anxiety and self-preoccupation, and many forms of socioemotional support. Given these beneficial side effects, participation in volunteer activities throughout adulthood should promote health. It should also contribute to activity in old age and, consequently, to multiple roles later in life. But again it is not clear how the duration, number of spells, or the timing of volunteer work might influence women's health.

Objective Social Conditions

Health in childhood and early adulthood as well as a high socioeconomic status are resources that should promote health later in life. Educated women are more likely than those with fewer years of schooling to adopt behavior conducive to health throughout the life course, and a high socioeconomic status should provide the economic means for optimal health care. Educational achievement and occupational prestige have been shown to be highly correlated with subjective health appraisals (Nathanson 1980). Moreover, life-styles incorporating risk factors such as smoking have been linked to social class, with working-class women slightly more likely to smoke than higher-status women (Nathanson and Lopez 1987).

Aging and the Life Course

The life-course perspective offers no simple interpretation of the relationship between social integration earlier in life and health in the later years of adulthood. Rather, it suggests that the connection be viewed as a dynamic, and possibly cumulative, process, wherein the roles individuals occupy in early adulthood play out in patterns of involvement over the life course that, in turn, sustain health. The life-course perspective also places women's social integration and health within the context of a changing society. For example, later cohorts of women may be healthier than earlier cohorts precisely because they are more socially integrated. This hypothesis is raised by Verbrugge (1983, pp. 170–71), who suggests that "as women engage in more social roles, they will be less frequently ill in middle and older ages. Their rates of acute conditions and short-term disability will decrease." Because health has been shown to be linked to multiple roles, it is important to assess factors affecting the number of roles occupied in later adulthood as well as health. House, Umberson, and Landis (1988) point out that almost no attention has been paid to the determinants of such social integration.

In this study we seek to better specify the corollaries of successful

aging, in the form of social integration and health, in women's lives. We address the following questions: (1) What is the relationship between social integration (multiple roles) in early adulthood and social integration and health in the later years? (2) Does the duration of particular roles over the life course, rather than the mere sum of roles, influence subsequent social integration and health? (3) Is the timing or sequencing of roles related to health and social integration in the later years of adulthood? (4) Are role strains and psychological well-being—in the form of life satisfaction and self-esteem in early adulthood—related to social integration and health later in life?

METHODS

We draw on data from a two-wave panel study of women who were interviewed in 1956 in an upstate New York community, and reinterviewed 30 years later, in 1986. A random sample of 427 women who were both wives and mothers were interviewed in 1956 (Dean and Williams 1956). This same group was located and reinterviewed 30 years later, in 1986. We were able to find 76.3% (326) who were still alive in 1986 and to ascertain that 19.2% (82) had died during the 30-year period. Only 4.4% (19) of the original sample could not be located. This "catch-up" design (Kessler and Greenberg 1981) of the Women's Roles Study offers the rare opportunity to examine the long-term association between multiple roles and health.

Of the 326 who were alive and contacted in 1986, only 3% (13) refused to be interviewed. We draw on information from the initial interview in 1956 as well as the second interview in 1986, when extensive life-history material was collected, in addition to responses to more structured questions.

Outcome Variables

We have four measures of health, all highly correlated but each tapping something quite distinctive. The first (HLTHDUR) represents the duration of time from 1956 until the onset of a serious disease or disability. The date of the first onset of a lasting debilitating illness/condition (such as a heart attack, stroke, cancer, arthritis) was recorded for each respondent, drawing on life-history data. Those women who had not yet experienced a major illness by 1986 were coded as censored. The second, a measure of functional ability in 1986 (ABILIT), is a scale developed by Rosow and Breslau (1966) to measure "functional health" (see also Stahl 1984). It consists of six items answered by yes-no and multiple-choice responses (Rosow and Breslau [1966] reported a coefficient of reproduc-

ibility of .91). A healthy respondent reports that she can do the following things without help: go out to a movie, to church or a meeting, or to visit friends; walk up and down stairs; walk half a mile; do heavy work around the house. The third measure, health appraisal (APPRAIS), is a respondent's subjective evaluation of her health in 1986, on a scale ranging from 1 (poor health) to 10 (best health). The question reads: "Think of a scale from 1 to 10 where 1 is a person with *very serious health problems* and 10 is a person in the *very best health*. Where would you place *YOURSELF*?" In the correlational analysis we include yet another, more objective measure of health, called health status (STATUS). It is based on data gathered from the life history of the respondent and represents a coder's rating of the presence of acute health symptoms or disability at the time of the 1986 interview. This is coded "1" if the respondent appears in good health or with only minor health problems and "0" if respondent suffers from a major illness or debilitating handicap. This measure is used to help appraise the validity of the other health measures, rather than as an outcome measure per se. The means and standard deviations of these measures are presented in the Appendix.

We also examined the social integration of women, in terms of their multiple-role involvements, in 1986 (ROLES86). Six roles are incorporated in this measure: worker, churchmember, friend, neighbor, relative, and club or organization member. We operationalize ROLES86 as the number of such positions actively occupied by women. Thus, it is the enactment of religious affiliation, by attending services on a regular basis, rather than mere church membership, that matters. However, we use membership in clubs or organizations rather than attendance since our analysis shows that such membership, unlike church membership, is virtually synonymous with attendance. This ROLES86 measure is similar to that employed by Thoits (1986), except that in this sample we exclude marriage and motherhood in this measure of roles, in order to keep it similar to the 1956 measure in which all the respondents were married and had children present in the home. (We do, however, include whether or not they were married or had children present in 1986 in the analysis.)

Independent Variables and Analysis Plan

We include a 1956 measure of multiple roles (ROLES56, similar to ROLES86, described above), as well as a measure of health status during and before 1956. We also include measures of the timing, duration, and number of spells in employment and volunteering as well as current marital status, current household composition, number of children, duration of marriage, duration of time spent caring for preschoolers, and duration of time spent caring for other relatives (ailing husbands, parents,

or retarded children). The operationalization of these variables is described in the Appendix.

We examine a series of nested multiple regression models to estimate the effects of multiple roles in 1956, as well as the duration and timing of particular roles throughout adulthood, on social integration in 1986 (ROLES86) and on two measures of health in 1986 (ABILIT, APPRAIS). We calculate the effects of number of roles in 1956 as well as the independent effects of particular roles, such as employment, on health and social integration in 1986. We also include composite measures of the duration of employment, volunteer work, caregiving, marriage, and caring for preschoolers for the period between ages 18 and 55. Number of spells and the timing of roles are also considered, with these measures derived from the life-history interview material.

We also examine the rate at which individuals who were healthy in 1956 ($N = 271$) experienced a severe illness or disability from that time until 1986 (HLTHDUR). Event-history techniques were used to estimate the duration of health (Tuma and Hannan 1984). These procedures are good with censored data, producing estimates that are asymptotically unbiased and normally distributed (Kalbfleisch and Prentice 1980). The covariates are fixed, drawn from the 1956 survey. We estimate a Weibull model of age dependence, presupposing time dependence, such that as women grow older they are more likely to experience illness or disability.

Specifically, the model considered for HLTHDUR is:

$$Y = \log(\text{AGE AT ILLNESS}) = a + b1\text{AGE} + b2\text{EDUC56} +$$

$$b3\text{NUMKIDS56} + b4\text{ROLES56} + W,$$

where W is a disturbance term with an extreme value distribution (Kalbfleisch and Prentice 1980).[4]

Using these panel data and dynamic event-history techniques tells us more than we could get from cross-sectional data examining multiple roles and health, since we are able to estimate not merely whether women are sick or disabled at any one point in time but how multiple roles earlier in life relate to the duration of their health as well as to their social integration and health in the later years of life.

RESULTS

The relationship between social integration and health can be seen in the correlations presented in table 1. Multiple-role occupancy in 1986 is

[4] We used PROC LIFEREG in SAS (1985) to obtain maximum-likelihood estimates from the censored data.

TABLE 1

CORRELATIONS OF HEALTH MEASURES IN 1986 AND NUMBER OF ROLES
OCCUPIED IN 1956 AND 1986

Variables	1	2	3	4	5	6
1. Roles (1956)	.					
2. Roles (1986)	.218**					
3 Functional ability (1986)	165**	.232**				
4. Health appraisal (1986)	.112†	.224**	.699**			
5. Health status (1986)	108†	133*	.393*	.508**		
6. Health duration (1956–1986)	.089	.123*	.390**	.487**	.434**	. . .

† $P \leq .10$.
* $P \leq 05$
** $P \leq .01$

positively related to all four measures of health in 1986. This is not
surprising, since those not healthy in 1986 would be less likely to be
engaged in multiple roles. But note that multiple-role occupancy in *1956*
also is positively correlated with three of the health outcomes 30 years
later, in 1986. This provides preliminary evidence supporting the notion
of the salutary effects of social integration in earlier adulthood for subse-
quent health. This relationship could, however, be a spurious one, with
ROLES56 being, in fact, merely a surrogate for health or socioeconomic
status earlier in life. Accordingly, we include measures of previous health
and social status in our multivariate analysis, to see whether or not
ROLES56 continues to be positively related to health in 1986, controlling
for health in and prior to 1956, as well as socioeconomic status.

Duration of Health

Health itself can be considered as following a trajectory, changing over
the course of adulthood. The life-history data provide a unique opportu-
nity to study the duration of time from 1956 until the onset of a serious
disease or disability. Over a third (36%) of the 271 women who were
healthy in 1956 experienced a serious illness before the interview date in
1986.[5] The first model tested, using maximum-likelihood techniques to
estimate the duration of health (HLTHDUR), is the baseline model
(model 1, table 2). As expected, we found the duration of health to be
significantly and negatively related to age. The older the respondent was
in 1956, the sooner she was likely to experience a severe illness. More-

[5] Forty-two women are eliminated from the analysis because of missing information
or illness prior to 1956.

1621

TABLE 2

ESTIMATES OF THE DURATION OF WOMEN'S HEALTH, 1956–1986

	MODEL		
VARIABLES	1	2	3
INTERCEPT	4.40	4.09	4.31
AGE	−.02†	− 02†	− 02*
EDUC56	−.01	−.02	−.06†
NUMKIDS56 ... .---- ..	03	03	.02
ROLES56ᵃ . ...·		08	01
CLUB/ORG56 . . .. ·			.35**
Scale parameter	54ᵇ	54ᵇ	54ᵇ
Log likelihood	− 223 62	− 222.39	− 220.28

NOTE.—N of spells = 271
ᵃ Includes all roles except that ROLES56 in model 3 does not include membership in clubs and organizations, which is measured by CLUB/ORG56.
ᵇ Scale factor is significantly less than unity
† $P \leq .10$.
* $P \leq 05$
** $P \leq 01$

over, the scale parameter is smaller than unity for all models, suggesting that the rate of illness increases over the time period studied, such that women in the sample become more prone to illness over the 30 years between interviews. Other control variables—educational level and number of children—were not significantly related to duration of health.

Neither ROLES56 nor its logarithm was significant in estimating the duration of health. And women's psychological well-being in 1956, as measured by satisfaction with life, self-esteem, and the degree of role strain, was also unrelated to health duration. However, in disaggregating ROLES56, we found that involvement in clubs and organizations in 1956 (CLUB/ORG56) had a significant and positive relationship with the duration of health. In other words, of the women who were healthy in and before 1956, those engaged in nonpaid club or organizational activities in 1956 were more likely to remain healthy longer (see model 3, table 2).[6]

---

[6] We also tested a monotonic transformation (log) of ROLES56, since multiple roles may well not have a linear relationship with duration of health; however, neither measure was significant. We anticipated that the duration of volunteer work, which is weakly related to the CLUB/ORG56 measure ($r = .16$, $P = .006$), may also be an important factor in promoting the duration of health, as might the timing of volunteering. However, these measures are confounded with the dependent variable in that the duration of volunteering may precede, overlap, or extend beyond the duration of health.

Functional Ability and Health Appraisal

The first models tested in the multiple regression analyses were the base-line models—whether measures of women's health in 1986 (ABILIT and APPRAIS) were related to age, educational level, earlier health, and current family status. As expected, both the measure of functional ability (ABILIT) and subjective health (APPRAIS) in 1986 were negatively related to previous poor health (in or before 1956) as well as to age (see model 1, tables 3 and 4). Socioeconomic status, as measured by women's education, was not related to either functional ability or subjective health assessment in 1986.

Several measures of current family status (married in 1986, living alone in 1986, having adult children at home in 1986, and total number of children in 1986) were included, one at a time, in the baseline model. These were not related to women's functional ability or the appraisal of their health in 1986. But, given their theoretical significance, we kept both marital status in 1986 and total number of children in the model (see model 1, tables 3 and 4).

Adding ROLES56 to the equations produced a positive effect on the two measures of health (ABILIT and APPRAIS) 30 years later, net of the effects of age, education, earlier health, and current family status (see model 2, tables 3 and 4). If multiple roles have positive health outcomes, then role accumulation over the 30-year period should prove beneficial and, correspondingly, role loss should be detrimental to health. We found that an increase in number of roles from 1956 to 1986, as compared with no role change, was positively related to women's health appraisal while a loss of roles was negatively related to activity level (see model 3, tables 3 and 4). But this could reflect the fact that poor health later in life leads to role loss while good health leads to role accumulation, a possibility we consider below. What is important to note in model 3 is that multiple-role involvement in 1956 (ROLES56) remains statistically significant and positively related to health in 1986, net of subsequent role gains or losses.

We next addressed the health implications of the duration of particular roles. We included the length of time from ages 18 to 55 that women were involved in five roles: employment, volunteering, caregiving, caring for preschoolers, and marriage. We found that duration of caregiving had a marginally significant (P = .10) and negative impact on both the subjective appraisal of health (APPRAIS) and functional ability (ABILIT). The duration of marriage was positively related to functional ability at the .09 level of significance (see model 4 in table 3). The number of spells in each role had no effect on any of our measures of health (not shown).

TABLE 3

REGRESSION ANALYSIS FOR HEALTH ACTIVITY IN 1986 (ABILIT)[a]

VARIABLES	MODEL								
	1	2	3	4	5	6	7	8	9
AGE	−.29**	−.30**	−.30**	−.29**	−.28***	−.28***	−.27***	−.28***	−.23**
POOR HEALTH56	−.17***	−.18***	−.18***	−.18***	−.17***	−.19***	−.17***	−.18***	−.08***
EDUC56	.04	−.02	−.04	−.02	−.04	−.05	−.06	−.09	−.06**
MARRIED IN 86	.03	.02	.01	−.04	.02	.00	.01	.00	−.01
TOTNUMKIDS	.02	.03	.02	.02	.01	.02	.02	.02	.00
ROLES56[b]	.21**	.21**	.26**	.21**	.22**	.21**	.22**	.11*	.10†
ROLE CHANGE 56–86 (vs. no role change):									
Role gain 56–86			.04						
Role loss 56–86			−.14†						
DUREMP 18–55				.04					
DURVOL 18–55				.05					
DURCAR 18–55				−.10†					
DURPRESC 18–55				.01					
DURMAR				.11†					

EMPLOYMENT TIMING

(vs. never employed).[c]									
Employed continuously				-.07					
Employed intermittently				-.16†					
Started employment after age 40				-.08					
Stopped employment before age 40				-.21**					
VOLUNTEER TIMING									
(vs. never volunteered).[c]									
Volunteered continuously					.03			.03	-.00
Volunteered intermittently					.16*			.15*	.13†
Started volunteering after age 40					.04			.04	.03
Stopped volunteering before age 40					.06			.05	.03
EMPEVER						-.12*			
VOLEVER							.10†		
CLUB/ORG56								.22**	.17**
ILL56–86								.16	-.40**
Adjusted R^2	.11	.14	.15	.16	.16	.15	.17	.16	.31
f-test	7.51***	8.60***	7.05***	5.63***	6.28***	5.69***	7.89***	5.78***	11.11***
N	277	277	271	277	277	277	277	277	277

[a] The numbers reported here are standardized regression coefficients
[b] Includes all roles except those considered separately in models 4–9. For example, in model 4 ROLES56 does not include employment, which is measured by DUREMP18–55
[c] Measured from respondent's first marriage until age 55
† $P \leq .10$.
* $P \leq .05$.
** $P \leq .01$
*** $P \leq .001$

TABLE 4

REGRESSION ANALYSIS FOR HEALTH APPRAISAL IN 1986 (APPRAIS)[a]

VARIABLES	MODEL								
	1	2	3	4	5	6	7	8	9
AGE	-.17**	-.17**	-.16*	-.16*	-.17**	.14*	-.16*	-.15*	-.10
POOR HEALTH56	-.12*	-.12*	-.12†	-.13*	-.11†	-.13*	-.11†	-.12*	.00
EDUC56	.02	-.02	-.03	-.02	-.03	-.04	-.05	-.07	-.03
MARRIED IN 86	.06	.06	.03	.00	.05	.05	.04	.05	.04
TOTNUMKIDS	.08	.09	.06	.15	.08	.09	.08	.09	-.06
ROLES56[b]		.15*	.21**	.15*	.15*	.15**	.15*	.08	.06
ROLE CHANGE 56-86 (vs. no role change):									
Role gain 56-86			.13†						
Role loss 56-86			-.05						
DUREMP 18-55				.08					
DURVOL 18-55				.05					
DURCAR 18-55				-.10†					
DURPRESC 18-55				.07					
DURMAR 18-55				.09					

1626

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9
EMPLOYMENT TIMING (vs. never employed):c									
Employed continuously					– 13*				
Employed intermittently					– 09				
Started employment after age 40					–.06				
Stopped employment before age 40					–.23**				
VOLUNTEER TIMING (vs. never volunteered) c									
Volunteered continuously						.02		.02	–.01
Volunteered intermittently						11		09	.08
Started volunteering after age 40						.00		.00	– 02
Stopped volunteering before age 40						10		09	.08
EMPEVER							–.09		
VOLEVER							07		
CLUB/ORG56								.17**	10†
ILL56–86				07	10	.06	.07	10†	– 49**
Adjusted R^2	04	.06	07	07	10	.06	.07	.07	.30
f-test	3.53***	4.08***	3 36***	2.83***	3.86***	2 78***	3 58***	2.85***	10.63***
N	275	275	270	275	275	275	275	275	275

[a] The numbers reported here are standardized regression coefficients.

[b] Includes all roles except those considered separately in models 4–9. For example, in model 4 ROLES56 does not include employment, which is measured by DUREMP18–55

[c] Measured from respondent's first marriage until age 55.

† $P \leq .10$

* $P \leq .05$

** $P \leq .01$

*** $P \leq .001$

It may be neither the length of time occupying roles nor spells in roles, but their *timing* that is most important for health. Accordingly, we examined the impacts of early, late, intermittent, or continuous involvement in two roles, those of worker and volunteer. The omitted category in each case was never having occupied the role. We found that both working for pay intermittently throughout adulthood and leaving the labor force early (before age 40) were negatively related to functional ability in 1986 (see model 5, table 3). Leaving the labor force early in adulthood was also negatively related to health appraisal in 1986 as was working continuously throughout adulthood (see model 5, table 4). Women who volunteered on an intermittent basis were likely to score higher on functional ability in 1986 than those who had not (see model 6, table 3).

Another consideration might be whether one had worked or volunteered at all, rather than the timing or duration of these roles. We included two dichotomous variables and found that ever having volunteered was positively related to functional ability in 1986 while ever having been employed was negatively related to this measure of health. Although the signs are in the same direction for health appraisal, neither variable was significant (see model 7, tables 3 and 4).

Measures of psychological well-being and role strain in 1956 were not related to either measure of health in 1986. In addition, we speculated that it may be the timing of marital dissolution, rather than marital status at a particular point in time that is related to health. However, we found that having one's marriage dissolve before or after age 40 had no impact on functional ability or subjective health appraisal in 1986 (results not shown).

Throughout all the models tested, social integration, as measured by multiple roles in 1956 (ROLES56), persisted as positively related to women's functional ability and to their subjective ratings of their health in 1986. In examining the independent effects of the six components of roles in 1956—worker, churchmember, friend, neighbor, relative, and club or organization member—we found that club or organizational involvement in 1956 had a significant, positive association with both functional ability and subjective health appraisal in 1986. To explore further the importance of club/organizational membership in 1956, we examined its relationship to the volunteer role. We found that CLUB/ORG56 was positively but weakly correlated to volunteering intermittently ($r = +.11, P = .06$), so these are not simply alternative measures of the same phenomenon. Moreover, both belonging to a club or organization in 1956 *and* volunteering intermittently from marriage to age 55 had significant and positive associations with women's functional abilities in 1986 (see model 8, table 3).

1628

Multiple roles in 1956 and, especially, club membership may in some way reflect socioeconomic status not captured by educational level. To test that possibility we included a number of alternative measures of socioeconomic status in our estimations. We found that, even when husband's occupation, social class of family of orientation, social class in 1956, or income in 1986 (as a surrogate for income in 1956 that was not available) are controlled for, ROLES56 continued to be significantly related to functional ability in 1986, and CLUB/ORG56 remained significant for both functional ability and health appraisal in 1986 (data not shown).

Multiple-role involvements, including club/organizational memberships, could also be a proxy for health, with healthy, active women more likely to be social participants. However, controlling for serious illness between 1956 and 1986 does not alter the relationships between these variables and our measures of health in 1986 (see model 9 in tables 3 and 4), suggesting that role affiliations are tapping more than physical well-being.

Social Integration, 1986

Another component of successful aging is active involvement in a number of social roles. We found that women occupying multiple roles in 1956 were the most likely to be occupying multiple roles in 1986, net of other background factors (see model 2, table 5). Duration of work and volunteer roles were also related to multiple roles in 1986, suggesting that extended involvement in nonfamily roles throughout adulthood promotes multiple-role occupancy later in life (see model 3, table 5). Indeed, women who had worked continuously from age 18 to 55 were more likely to have multiple roles in 1986 (see model 4, table 5). Any involvement as a volunteer, regardless of when it occurred, was positively related to multiple roles in later adulthood (see models 5 and 6, table 5). In addition, women with larger families were more likely to occupy multiple roles in 1986. Psychological well-being in 1956 was unrelated to multiple-role occupancy 30 years later, in 1986 (data not shown). Controlling for serious illness between 1956 and 1986 (ILL56–86) does not affect the positive relationship between volunteering at any time during adulthood or the number of roles occupied in 1956, on the one hand, and occupying multiple roles in 1986 on the other.

To see the relationship between earlier role configurations and multiple roles in old age we ran analyses for those women who, in 1986, ranged in age from 64 to 81, past the typical retirement age for women in the United States. For this group the duration of paid work and volunteer work in the years from ages 18 to 55 remained important predictors of

TABLE 5

REGRESSION ANALYSIS FOR NUMBER OF ROLES, 1986 (ROLES86)[a]

VARIABLES	MODEL						
	1	2	3	4	5	6	7
AGE	− .10	− .12†	− .10†	− .08	− .08	− .04	− .07
POOR HEALTH56	− .10†	− .11†	− .02	− .12*	− .12*	− .12*	− .10†
EDUC56	.12*	.07	.02	.09	− .01	.02	.00
MARRIED IN 86	− .02	− .02	− .06	− .02	− .06	− .06	− .06
TOTNUMKIDS	.12*	.13*	.01	.14*	.13*	.12*	.13*
ROLES56[b]		.23**	.17**	.21**	.21**	.18**	.20**
DUREMP 18–55			.16*				
DURVOL 18–55			.28†*				
DURCAR 18–55			.02				
DURPRESC 18–55			.16				
DURMAR 18–55			.02				
EMPLOYMENT TIMING (vs never employed)[c]							
Employed continuously				.06*			
Employed intermittently				.13			
Started employment after age 40				.04			
Stopped employment before age 40				.10			

1630

VOLUNTEER TIMING							
(vs. never volunteered):c							
Volunteered continuously					.18**		.17**
Volunteered intermittently					33**		.32**
Started volunteering after age 40					13*		13*
Stopped volunteering before age 40					15*		.15*
EMPEVER						08	
VOLEVER						.23**	
ILL56–86							−.08
Adjusted R^2	,04	.08	15	.07	.13	.13	.14
f-test	3.29***	5.29***	5.57***	3.17***	5 39***	6.22***	5.11***

social integration in their later years, as did number of roles in 1956. And, even for these older women, any involvement—early, late, intermittent, or continuous—as a volunteer served to promote multiple-role occupancy in old age (data not shown).

DISCUSSION AND CONCLUSIONS

We have tried to add to the understanding of the links between social integration and women's health by adopting a more dynamic, life-course perspective, examining the timing, duration, and spells of role involvement and looking at social integration (in the form of multiple roles) as both an outcome and a predictor variable, at two different points in time, spanning a 30-year interval. Social integration, as we operationalized it, consists of involvement in multiple roles beyond those of wife and mother. Such integration serves to prevent the social isolation that for some may be concomitant with aging.

We found that occupying multiple roles in 1956 was positively associated with women's health and their social integration in 1986, net of a range of background variables, including age and previous health. There is, most likely, nothing unique about the year, 1956, in which these women were interviewed. Rather, the multiple roles women occupied in that survey year can be viewed as an instrumental variable, depicting the active role involvements of women, possibly throughout early and middle adulthood. From this point of view, social integration during the adult years appears to promote both social integration *and* health in the later years of women's lives.

One role in particular, being a member of a club or organization, appeared to be especially salutary, again after previous health is controlled for. This remained the case even when a number of objective and subjective measures of social class were controlled, suggesting that CLUB/ORG56 is more than a surrogate for socioeconomic status.

Paid work over the life course, while positively related to multiple roles later in life, was negatively related to our measures of health. On the other hand, volunteer work on and off through adulthood was positively related to health. The duration of time spent caring for infirm or aged parents, spouses, or other relatives was negatively related to health in 1986.

The pathways to successful aging appear to be grounded in initial good health. Women with health problems during childhood or prior to the 1956 interview were likely to report less functional ability and a lower health appraisal in 1986 than were those without early problems. This

supports a cumulative, accentuation process, in which the least healthy are the most vulnerable later in life.

A similar process seems to operate in terms of social integration, with multiple-role occupancy in 1956 portending multiple-role occupancy in 1986, and the duration of employment or volunteer roles promoting continued multiple-role involvement. However, any voluntary participation at any time in adulthood appears to promote multiple-role occupancy in the later years.

House, Umberson, and Landis (1988) suggest that sociologists have much to contribute to the study of "the existence, structure, and content of social relationships as they relate to health" (p. 315). What is the broader significance of the findings produced in this study? First, what Coser (1975) refers to as a "plurality of life worlds" in the prime of adulthood appears to foster physical health, as well as multiple-role involvements, later in life. Second, social participation, as members of clubs or organizations or as unpaid, volunteer workers, seems especially conducive to subsequent health and integration. Whether this reflects women's active choices, the personality characteristics related to being "joiners," the social contacts, support, and other psychosocial payoffs of social participation, or some other, unmeasured construct remains to be investigated.

Third, these findings highlight the fact that individuals traverse two distinctive tracks throughout their lives: (1) life-course pathways consisting of various role transitions and trajectories and (2) developmental pathways, including trajectories and transitions related to physiological health. These processes are closely related, but not in any simplistic or additive fashion. For example, participation in volunteer work, intermittently or at any time in adulthood, not the duration of volunteering, is what seems to matter for successful aging. And labor-force participation, while conducive to multiple-role occupancy later in life, does not appear to promote subsequent health.

"Successful aging" can be depicted as living *both* healthy and active, involved lives. Both social integration and health in the later years may reflect choices and experiences throughout adulthood. But note that this sample consists of women who were wives and mothers in 1956, who were therefore socially integrated to some degree by virtue of these family roles. Nonmarried women may experience nonfamily roles differently from the women we have studed here. However, by 1986 only 65% of our sample remained married; thus the protective function of marriage was no longer available for a considerable minority of these women. For the women in our sample, all of whom had at one time been married and been mothers (which reflects the experience of the vast majority of

women in the United States), our findings suggest that the aging process may vary widely as a result of other social roles in the prime of life. The timing and number of spells of particular roles appear to be less important than multiple-role involvements per se. But the role of social participant, as a volunteer at any time during adulthood or as a club member, seems to influence the ways in which women, at least those who have been wives and mothers, grow old.

Still, the mechanisms by which social integration promotes good health remain open to speculation. Socially active women may be involved in community activities despite their family responsibilities because of their high levels of vitality and energy, sense of civic responsibility toward others, and/or their socially gregarious personalities. Thus, the sociological notion of social integration may be augmented by a biopsychological interpretation: women with multiple roles, those who volunteer and who join clubs and organizations, may possess a high degree of zest and vitality on the biological side, and a gregarious social nature on the personality side. If we were to design a prospective, instead of follow-up, study of successful aging we would want to examine more closely the biological, psychological, social, and structural antecedents of multiple-role involvements generally and social participation in volunteer work, clubs and organizations specifically. We also would want to ascertain the processes and circumstances giving rise to role acquisition and loss, including the subjective interpretations of these transitions, how role attachments evolve over time, and their various consequences for women's (and men's) lives.

Certainly, occupying multiple roles within society augments an individual's social network, power, prestige, resources, and emotional gratifications. Bronfenbrenner (1979, p. 6) hypothesizes that an ever-broadening role repertoire promotes human development, in part by facilitating interaction with persons occupying a variety of roles. He notes that "roles have a magic-like power to alter how a person is treated, how she acts, what she does, and thereby even what she thinks and feels." But particular roles can also be detrimental to health, as we found in the case of caregiving. This suggests that preference and choice, as well as level of autonomy, may be important elements in linking role attachments to health (see also Coser 1975). And satisfaction in a role, rather than simply role occupancy, may be key. For example, volunteering and club memberships are more discretionary than caregiving or even employment, and we found both organizational membership and the unpaid volunteer role to be positively associated with women's health. Social researchers must attend not only to social integration, but the degree to which roles operate as constraints versus options and opportunities throughout the course of women's lives.

1634

APPENDIX

TABLE A1

Variable Definitions and Distributions

Variables	Definition	Frequency (%)	Mean	SD
ABILIT	Functional ability in 1986, sum of six yes-no and multiple-choice items· 0 = worst health, 6 = best health (Rosow & Breslow 1966)		4 09	1.73
APPRAIS	Health appraisal in 1986, scale from 1 to 10, 10 = very best health		7 09	2.29
STATUS	Coders' rating of health status in 1986, 1 = good health	89		
HLTHDUR	Duration of time 1956 until onset of serious disease/disability or 1986		20.14	8.35
ROLES86	Summed number of role attachments in 1986, includes memberships in clubs or organizations, employment status, religious attendance, seeing relatives, visiting with neighbors, having one or more close friends		3.55	1 23
AGE	Age (year to nearest decimal) in 1956		35 27	6.70
EDUC	Educational level in 1956, 1 = 8th grade, 6 = advanced study		3 33	1.09
POOR HEALTH56	1 = ill in 1956 or before	9		
ILL	Serious illness between 1956 and 1986, 1 = ill	40		
MARRIED86	1 = married	65		
TOTNUMKIDS	Number of children		3 24	1.69
ROLES56	Summed number of role attachments in 1956 (see ROLES86)		4.20	0 99
DUREMP	Number of years employed (from ages 18 to 55)		16 24	9.93
DURVOL	Number of years volunteering (from ages 18 to 55)		12.40	11.80
DURCAR	Number of years caring for others (from ages 18 to 55)		2 32	4.65
DURPRESC	Number of years caring for own preschoolers (from ages 18 to 55)		12.75	4.69
DURMAR	Number of years married from ages 18 to 55		30.85	5 37

Note —A more complete description of the variables is available from the authors

TABLE A2

Role Trajectory Variables: Definitions and Distribution

Variables	Frequency (%)
Employment timing from marriage to age 55:	
Never employed	11.6
Employed continuously	2.9
Employed intermittently	66 9
Started employment after age 40	6.1
Stopped employment before age 40	12 5
Volunteer timing from marriage to age 55.	
Never volunteered	24.4
Volunteered continuously	7.9
Volunteered intermittently	44.3
Started volunteering after age 40	9.3
Stopped volunteering before age 40	13.7
Role change, 1956–86·	
Role gain from 1956–86	20
Role loss from 1956–86	56
No change in number of roles	24

Sources —1956 Cornell Study of Women's Roles, National Institute of Mental Health M-902, principal co-investigators: John P. Dean and Robin M. Williams, Jr., and the 1986 Women's Roles Project, National Institute of Aging R01 AGO 5450, principal co-investigators Phyllis Moen and Robin M. Williams, Jr.

REFERENCES

Aneshensel, Carol S., Ralph R. Frerichs, and George, J. Huba. 1984. "Depression and Physical Illness: A Multiwave, Nonrecursive Causal Model." *Journal of Health and Social Behavior* 25·350–71.

Antonovsky, Aaron. 1987. *Unraveling the Mystery of Health: How People Manage Stress and Stay Well.* San Francisco: Jossey-Bass.

Berkman, Lisa F., and Lester Breslow. 1983. *Health and Ways of Living: The Alameda County Study.* New York: Oxford University Press.

Bronfenbrenner, Urie. 1979. *The Ecology of Human Development.* Cambridge, Mass.: Harvard University Press.

Coser, Rose. L. 1975. "The Complexity of Roles as a Seedbed of Individual Autonomy." Pp. 237–64 in *The Idea of Social Structure: Papers in Honor of Robert Merton,* edited by L. A. Coser. New York: Harcourt Brace Jovanovich.

Dean, John P., and Robin M. Williams, Jr. 1956. "Social and Cultural Factors Affecting Role-Conflict and Adjustment among American Women: A Pilot Investigation." Progress report submitted to the National Institute of Mental Health, Bethesda, Md.

Durkheim, Emile. (1897) 1951. *Suicide: A Study in Sociology,* translated by John A. Spaulding and George Simpson. Glencoe, Ill.: Free Press.

Elder, Glen H., Jr. 1985. "Perspectives on the Life Course." Pp. 23–50 in *Life Course Dynamics: Trajectories and Transitions, 1968–1980,* edited by Glen H. Elder, Jr. Ithaca, N.Y.: Cornell University Press.

Fredrickson, Barbara L., and Laura L. Carstensen. 1990. "Choosing Social Partners:

How Old Age and Anticipated Endings Make People More Selective." *Psychology and Aging* 5:335–47.

Gove, Walter R , and M. Hughes 1979. "Possible Causes of the Apparent Sex Differences in Physical Health: An Empirical Investigation." *American Sociological Review* 44:126–46.

House, James S., Cynthia A. Robbins, and Helen L. Metzner. 1982. "The Association of Social Relationships and Activities with Mortality: Prospective Evidence from the Tecumseh Community Health Study." *American Journal of Epidemiology* 116:123–40

House, James S., Karl R. Landis, and Debra Umberson. 1988. "Social Relationships and Health." *Science* 241.540–45.

House, James S., Debra Umberson, and Karl R. Landis. 1988. "Structures and Processes of Social Support " *Annual Review of Sociology* 14:293–318.

Hughes, Michael D., and Walter R. Gove 1981. "Living Alone, Social Integration, and Mental Health." *American Journal of Sociology* 87:48–74.

Kalbfleisch, John D., and Ross L. Prentice. 1980. *The Statistical Analysis of Failure Time Data*. New York: Wiley.

Kessler, Ronald C., and David F. Greenberg. 1981. *Linear Panel Analysis: Models of Quantitative Change*. New York: Academic Press.

Kobrin, Francis, and Gerry Hendershot. 1977. "Do Family Ties Reduce Mortality? Evidence from the United States, 1966–1968." *Journal of Marriage and the Family* 39:737–45.

Litwak, Eugene. 1985. *Helping the Elderly: The Complementary Roles of Informal Networks and Formal Systems*. New York: Guilford.

Moen, Phyllis, Donna Dempster-McClain, and Robin M. Williams, Jr. 1989. "Social Integration and Longevity: An Event History Analysis of Women's Roles and Resilience." *American Sociological Review* 54:635–47.

Morgan, David L. 1988. "Age Differences in Social Network Participation." *Journal of Gerontology, Social Sciences* 43 S129–S137.

Nathanson, Constance 1980. "Social Roles and Health Status among Women The Significance of Employment." *Social Science and Medicine* 14:463–71.

Nathanson, Constance, and Gerda Lopez. 1987 "The Future of Sex Mortality Differentials in Industrialized Countries: A Structural Hypothesis." In *Population Research and Policy Review* 6:123–36.

Riley, Matilda W , and Kathleen Bond. 1983. "Beyond Ageism: Postponing the Onset of Disability." Pp. 243–52 in *Aging in Society: Selected Review of Recent Research*, edited by Matilda White Riley, Beth B Hess, and Kathleen Bond. Hillsdale, N.J.: Erlbaum.

Riley, Matilda W., and John W. Riley, Jr. 1989. "The Lives of Old People and Changing Social Roles " *The Annals of the American Academy of Political and Social Science* 503:14–28.

Rosow, I., and N. Breslau. 1966. "A Guttman Health Scale for the Aged." *Journal of Gerontology* 21·556–59.

Rowe, John W., and Robert L. Kahn 1987 "Human Aging Usual and Successful." *Science* 237 143–49.

SAS Institute. 1985. *SAS User's Guide: Statistics*. Cary, N.C.: SAS Institute.

Sieber, Sam D. 1974 "Toward a Theory of Role Accumulation." *American Sociological Review* 39.567–78.

Smith, Ken R., and Phyllis Moen. 1988. "Passage through Midlife: Women's Changing Family Roles and Economic Well-Being." *Sociological Quarterly* 29:503–24.

Stahl, Sidney M. 1984. "Health." Pp. 85–116 in *Research Instruments in Social Gerontology*. Vol. 3, *Health, Program Evaluation and Demography* Edited by David J. Mangen and Warren A. Peterson. Minneapolis. University of Minnesota Press.

Thoits, Peggy A. 1983. "Multiple Identities and Psychological Well-Being. A Re-formulation and Test of the Social Isolation Hypothesis." *American Sociological Review* 48:174–87.

———. 1986. "Multiple Identities: Examining Gender and Marital Status Differences in Distress." *American Sociological Review* 51.259–72.

Tuma, Nancy B., and Michael Hannan 1984. *Social Dynamics: Models and Methods.* Orlando, Fla.. Academic Press

Umberson, Debra. 1987 "Family Status and Health Behaviors. Social Control as a Dimension of Social Integration." *Journal of Health and Social Behavior* 28:306–19.

Verbrugge, Lois M. 1983. "Multiple Roles and Physical Health of Women and Men." *Journal of Health and Social Behavior* 24.16–30.

———. 1989. "The Twain Meet Empirical Explanations of Sex Differences in Health and Mortality." *Journal of Health and Social Behavior* 30:282–304.

Waldron, Ingrid, and Jerry A. Jacobs. 1989. "Effects of Multiple Roles on Women's Health—Evidence from a National Longitudinal Study." *Women and Health* 15:3–19.

Shadow Education and Allocation in Formal Schooling: Transition to University in Japan[1]

David Lee Stevenson
U.S. Department of Education

David P. Baker
Catholic University of America

Shadow education is a set of educational activities that occur out-
side formal schooling and are designed to enhance the student's
formal school career. Analyses of data from a longitudinal study of
high school seniors in Japan indicate that students from higher
socioeconomic backgrounds are more likely to participate in shadow
education and that students who participate in certain forms of
shadow education are more likely to attend university. Expanding
theories of allocation to incorporate shadow education may enhance
the study of how students are allocated to places in formal schooling
and how social advantages are transferred across generations.

The capacity of formal schooling to differentiate students is central to
allocation theories of education (Bourdieu and Passeron 1977; Bowles
and Gintis 1976). Allocation processes, however, vary across educational
systems in terms of the criteria by which allocation decisions are made,
in the timing of such decisions, and in the factors that influence those
decisions (Eckstein and Noah 1989). Allocation processes also can have
institutional effects on schooling itself (Meyer 1977).

An example of such an institutional effect is the way allocation pro-

cesses encourage the development of shadow education, a set of educational activities outside formal schooling that are designed to improve a student's chances of successfully moving through the allocation process. We will examine the general characteristics of shadow education and describe its role in the transition from high school to university in Japan.

SHADOW EDUCATION

Certain characteristics of the educational allocation process foster the development of shadow education. First, and most important, is the use of formal examinations—particularly centrally administered examinations. Second, shadow education flourishes if schooling uses "contest rules" instead of "sponsorship rules" (Turner 1960). And third, shadow education is prevalent when there are tight linkages between the outcomes of educational allocation in elementary and secondary schooling and future educational opportunities, occupations, or general social status.

When these qualities occur together, extensive shadow education can be observed. For example, in Taiwan, graduates of elite universities have significant advantages in the labor market, and there is extensive shadow education to prepare students for university entrance examinations (Lin 1983). In Hong Kong, graduates' hiring and pay rates are partly dependent on their performance on secondary-school-certificate examinations, and shadow education takes the form of tutoring and after-school classes to prepare students for the examinations (Mitchell 1968; Sweeting 1983). In Greece, because admission to an elite secondary school is considered critical for acceptance at a prestigious university, students use tutors and attend special after-school classes to prepare for the national entrance examination for secondary school (Katsillis and Rubinson 1990).

The timing, use, and forms of shadow education are shaped by allocation rules. We use the term "shadow" to denote the strong connection between allocation rules and nonformal schooling; we do not imply that these activities are hidden. Indeed, in systems such as Japan's, these activities make up a large, open enterprise.

SHADOW EDUCATION AND CONTESTED SPONSORSHIP IN JAPAN

Educational allocation in Japan is organized around the institutional qualities that foster the development of shadow education. First and foremost is the use of formal examinations in the educational allocation process, particularly since the examinations are central to the allocation process, the weight of nonexamination criteria in the selection process is minimal, and there are few critical school-to-career transition points

(U.S. Department of Education 1987; Amano 1990). Second is the clear and immutable connections between the outcomes of the educational allocation process and future occupational positions or general social status (Brinton 1988; Rosenbaum and Kariya 1989). The Japanese combination of meritocratic examinations and rich educational rewards shapes an allocation process that can be described as "contested sponsorship."[2] Contested sponsorship arises from the organization of Japanese education that shifts from more compulsory and nonselective schooling in the lower grades to increasing levels of selectivity by the end of high school.

Compulsory primary education in Japan is organized along egalitarian policies with all students receiving the same curriculum at the same pace regardless of ability (Cummings 1980). Secondary education is more stratified than primary education because high schools are ranked by prestige (mostly determined by how well their students do on university entrance examinations; Amano 1986). Nevertheless a standard secondary curriculum is taught to all students and high schools differ chiefly in the degree of emphasis on preparation for the university entrance examination (Rohlen 1983; Shimahara 1984).

This period of relatively homogeneous education ends with high school graduation. The system of higher education in Japan is a clear, rigid hierarchy of educational opportunities controlled by meritocratic entrance examinations. All high school graduates are eligible to apply and a student's performance on the entrance examination is the key determinant of admission. Other criteria, such as high school grades, teachers' recommendations, extracurricular activities, and "character," are not important factors in the admission process since they are seen as "arbitrary" and "subjective" (Shimahara 1984). Once admitted most university students continue to study at the same university until graduation. Eighty-seven percent of university students eventually graduate, with 75% of all students graduating on schedule (Amano 1986).

Increasingly larger numbers of students attempt to enter college, in part because the potential prize is entrance to one of the most prestigious universities, which leads to recruitment into a sponsored elite.[3] As else-

---

[2] Brinton (1988) characterized Japanese education as a series of "sponsored contests" in which the parents support their children to compete. We used the term "contested sponsorship" to describe the nature of the transition from high school to university. Students, whether or not they are supported by their parents, can compete for a place at a prestigious university Once admitted, students become "sponsored" in the sense that they enter a socially recognized elite with known advantages. For such students completing college is neither rigorous nor competitive and upon graduation they are actively recruited by the prestigious companies and civil service departments.

[3] The number of students attending university has increased during the past three decades In 1950, approximately 10% of the college-age students were attending university and by 1982 this percentage had increased to over 36% (Amano 1986). The

where, universities in Japan have social charters: the ability "to define people as graduates and as therefore possessing distinctive rights and capacities in society" (Meyer 1970, p. 59). But the clear hierarchy of Japanese universities has produced a set of "narrow" charters so that *which* university one attends affects one's occupational career and general social status (Amano 1986; U.S. Department of Education 1987).

University graduates are recruited by firms directly from university and many spend their entire work career with their first employer (Brinton 1988). Recruitment by employers begins in November and, within a five-day period, almost 80% of the spring graduates have selected an employer (Ushiogi 1986). Major companies and elite departments of the civil service only recruit graduates of certain departments in the most prestigious universities. Graduates employed by major companies have life-long advantages over graduates employed by small firms in terms of salaries, promotion opportunities, benefits, job stability, and yearly bonuses (Ushiogi 1986).

The tight linkages between education and work and the narrow charters of Japanese universities produce intense competition for admission to the most prestigious universities. In 1980, for example, 13,000 applicants competed for 3,077 places at Tokyo University (Rohlen 1983). For admission to the prestigious departments of law, economics, and medicine, the ratio of applicants to places may reach 7:1.

It is clear, then, why Japanese students with college aspirations arduously prepare for the entrance examinations. The importance of this transition point is heavily emphasized in the name "examination hell" (*juken jigoku*). The intensity of the competition is captured in the popular phrase, "four pass, five fail," which admonishes students that those who sleep four hours pass their entrance examinations but those who sleep five hours fail.

The contested sponsorship process in Japan is ripe for the development of shadow education. The dominance of examinations, the social and financial rewards for success, and the concentration of allocation at one time point intensify the growth of shadow education.

There are two sets of Japanese shadow-education activities: one used primarily during secondary school and the other used immediately after high school. During the secondary school years, students participate in

---

increase in enrollments is primarily due to the expansion of the private sector of higher education, and almost three-quarters of university students attended private universities in 1983 (Kitamura 1986). Approximately one-third of students attending postsecondary schools attend two-year junior colleges. Eighty-four percent of these schools are private institutions, and almost 90% of the students attending are females. Less than 4% of the students who attend junior colleges transfer to four-year universities (Amano 1986)

after-school and weekend preparation activities such as private cram schools, correspondence courses, and practice examinations. If high school graduates do not earn admission to the university of their choice, they may spend one or more additional years preparing. Such students are known as *ronin,* a word that was used to describe lordless samurai. Ronin do not legitimately belong to either the world of formal schooling or the world of work. They spend their time engaged in preparation for the university entrance examinations, often in intensive preparation schools (*yobiko*).[4]

Shadow education encompasses a large set of varied educational activities that are firmly rooted within the private sector. Students and their families pay tuition for private schools to prepare them for examinations, purchase workbooks with questions from previous examinations, and pay for practice tests that are administered and graded by private companies. Since there are millions of students competing for university entrance examinations, shadow education is a vast market representing an industry worth 870 billion yen (in U.S. currency, approximately $7 billion) in 1986 (Katsuki 1988).

Despite the size of shadow education in Japan and in other educational systems, little is known about its role in the allocation of students. Which students participate in shadow education? What effects does participation in shadow education have on students' educational careers? Is shadow education an avenue for the transmission of social advantages from parents to their children in the contest for educational credentials? To answer these questions we analyzed a longitudinal data set from a study made of Japanese high school seniors at the time of their transition to postsecondary education or to work.

METHODS

The data come from a large longitudinal study of high school seniors in Japan conducted by the Youth Research Institute of Tokyo in 1980 and 1982 (Nihon Seishounen Kenkyuujo 1981; Nihon Seishounen Kenkyuujo 1984). The sample of 1980 seniors (base-year sample) is representative of high school seniors in Japan and includes 7,240 students, 4,280 of whom were reinterviewed in 1982 (follow-up sample) (Nihon Seishounen Kenkyuujo 1984). A multistage, probability-proportionate-to-size sample was chosen from selected geographical locations and then three to six high schools were randomly selected from each location. In each school two classrooms were selected and all students in the selected classroom partic-

---

[4] There was a similar status of intensive preparation for aspiring feudal Chinese literati (*shen-shih*) whose existence was referred to as an "examination life" (Franke 1963)

ipated in the study. The mix of schools (public vs. private) and the curriculum (academic vs. vocational) are proportional to national distributions.

MEASURES

Shadow Education

Students in the base-year sample were asked about their participation in different types of shadow education during high school and whether they would consider becoming a ronin after high school. The follow-up sample was asked whether they had became a ronin after graduating from high school. The main types of shadow-education activities high school students can use are:

1. Practice examinations (*mogi shiken*), provided and graded by private firms, assess a student's chances of being admitted to university. Students receive a report comparing their performance with national norms, are notified of subject areas that require greater study, and are given an estimate of their chances of being admitted to a particular university.

2. Correspondence courses (*tsushin tensaku*), purchased from mail-order companies, provide exercises for entrance examinations that are mailed back and then returned graded.

3. Private tutors (*katei kyoshi*) are primarily used for staying abreast of regular schoolwork and used less for examination preparation.

4. Private after-school classes (*juku*) come in two types: *gakushu juku*, which are remedial classes, and *shingaku juku*, for preparation for the university entrance examinations. *Juku* vary in the size of the classes from small classes that may meet in the home of the tutor to large classes in a school. Students in primary school may participate in *juku* to develop academic abilities in coordination with the curriculum of the formal school, while secondary students in *juku* prepare for university entrance examinations. There is a great deal of variation in the atmosphere and demands of *juku* but there is a general opinion that participation in a *juku* is an important part of a student's school career (Cummings 1980; White 1987).

5. Full-time preparation following high school (ronin) is a strategy used by students who do not gain admission to any university or by those who do not gain admission to the university they wish to attend. Ronin take a year or more to prepare solely for

university entrance examinations and often attend *yobiko,* a private examination preparation school (Tsukada 1988).

Other Measures

Two measures of school characteristics include (1) whether the student was taking vocational or academic courses in high school and (2) the student's rating of the relative reputation of his or her high school for educational attainment of past students. Both measures tap information that is well-known by Japanese students (Rohlen 1983). Additional information was collected from the students about the educational attainment of their parents, their family's income, their school performance, and their educational plans.[5] And in the follow-up, students were asked whether they had attended university or participated as ronins during the two years after high school.

RESULTS

Three specific questions about shadow education in Japan are explored: (1) How pervasive is participation in shadow education? (2) What are the determinants of participation in various forms of shadow education? and (3) What are the consequences of participation in shadow education for the educational allocation of students?

Prevalence of Shadow Education

Because shadow education is both a large industry and is publicly defined as useful for educational advancement in Japan, we would expect to find high proportions of students using these educational practices, particularly among those students who definitely plan to attend a university. Table 1 presents the percentage of students who participate in shadow-education activities during high school and the percentage who are willing to become ronin after high school. The first column shows the participation of all base-year students and the second column shows the participation of only those base-year students who have college plans (about 74% of the base-year sample).

Japanese students are voracious consumers of shadow-education activities. Eighty-eight percent of those students with college plans participated in at least one activity during high school and 60% participated in

[5] The family-background variables of parents' education, occupation, and income behave similarly in models of status attainment in Japan as in the United States (Fujita 1978).

TABLE 1

PARTICIPATION IN SHADOW EDUCATION BY JAPANESE STUDENTS DURING HIGH SCHOOL (BASE-YEAR SAMPLE)

SHADOW-EDUCATION ACTIVITY	TOTAL SAMPLE OF STUDENTS (%)	STUDENTS WITH COLLEGE PLANS[a] (%)	SCHOOL PERIOD WHEN STUDENT BEGAN TO PLAN FOR COLLEGE				
			Early Elementary School (%)	Late Elementary School (%)	Junior High School (%)	1st–2nd Year, High School (%)	χ^2
Practice examination (*mogi shiken*)	54	68	72	69	68	60	40 0*
Correspondence course (*tsushin tensaku*) . .	30	43	45	41	35	26	99.2*
Private tutor (*katei kyoshi*)	8	11	15	12	8	7	66.9*
After-school class (*juku*)	35	46	52	43	38	29	144.4+
Plans to be a ronin after high school	29	32	41	33	26	19	123.0*
N	7,240	5,352	1,740	826	1,924	862	

NOTE —Fifty-four percent of all students took a practice examination; 46% did not.

[a] Students with college plans represent 74% of base-year sample.

* $P < .00001$, $df = 3$.

two or more of these activities. Almost a third of the seniors were willing to become ronin after graduating if they needed additional preparation for the university examinations.

We have argued that the use of shadow education is related to its perceived educational benefit. If that is so, we should find that students' participation varies with their plans for postsecondary education. The last four columns in table 1 show that participation is related to the age at which students stated they planned to attend a university. Students who had college plans early in elementary school were most likely to use shadow education, while the rates of participation declined among late planners.

These findings are similar to other assessments of the pervasiveness of Japanese shadow education, which find that 40% of children in grades 6–8 in urban areas attend after-school classes and that one out of 10 high school students attend special preparation classes for university entrance examinations (Cummings 1980; Kondo 1974; Rohlen 1980, 1983; Tsukada 1986). In short, shadow education is institutionalized as a part of the Japanese educational allocation process.

Determinants of Participation in Shadow Education

Although shadow education is pervasive, participation may not be equally spread across all students. The transition to university offers significant benefits, and many students engage in extraordinary preparation activities, but, because these educational activities are expensive and time-consuming, we expect children from families with more resources to participate more in shadow education than children from poorer families. Differential investments in shadow education also can occur within families. For example, because of the better job prospects for Japanese males and perceptions about gender and social opportunity (Brinton 1988, 1989), families may tend to invest their resources in the education of sons rather than daughters. Or families may invest in shadow education for only their most promising child or do so only when they live in areas, such as cities, in which these activities are most available.

To examine this we estimated logistic regression equations in which the dependent variable is whether or not the student ever participated in a particular shadow-education activity. The independent variables are (1) indicators of the student's family's resources (father's education, mother's education, and family income); (2) student's gender; (3) student's academic standing; (4) the reputation of the student's high school; (5) student's curriculum track; and (6) if the student lived in an urban area. Estimates of these equations, shown in table 2, are based on base-

TABLE 2

INFLUENCE OF STUDENT, FAMILY, AND SCHOOL CHARACTERISTICS ON PARTICIPATION IN SHADOW EDUCATION OF STUDENTS WITH COLLEGE PLANS (BASE-YEAR SAMPLE)

INDEPENDENT VARIABLES	SHADOW EDUCATION				
	Practice Exam (Mogi Shiken)	Correspondence Course (Tsushin Tensaku)	Private Tutor (Katei Kyoshi)	After-School Class (Juku)	Plans to be a Ronin
Background factors.					
Gender (1 = male)	.74†	.45†	−.35†	.07	2.22†
	(.11)	(.08)	(.11)	(.08)	(.12)
Academic standing	.27†	.15†	.17†	.05†	.14†
	(.03)	(.03)	(.04)	(.02)	(.03)
Father's education	.19†	.12†	.09*	.16†	.21†
	(.05)	(.03)	(.04)	(.03)	(.04)
Mother's education	.19†	.13†	.25†	.16†	.07
	(.07)	(.04)	(.06)	(.04)	(.05)
Family income	.06	.04	.23†	.17†	.05
	(.04)	(.20)	(.04)	(.03)	(.03)
Curriculum (1 = academic)	2.64†	1.77†	.05	.73‖	1.65†
	(.16)	(.20)	(.24)	(.15)	(.22)
High school's reputation	.41†	.23†	.09	−.04	.20†
	(.05)	(.04)	(.06)	(.04)	(.04)
Rural/urban (1 = urban)	.35†	.11	.34†	.75†	.98†
	(.12)	(.08)	(.12)	(.08)	(.09)
Intercept	−.18	−1.87	−4.51	−2.67	−4.24
	(.32)	(.28)	(.40)	(.26)	(.33)
Model χ² compared to model with intercept only	616.89	329.86	146.07	355.64	829.06
df	8	8	8	8	8
N	3,051	3,033	3,018	3,053	3,069

NOTE.—Entries are maximum-likelihood parameters from logistic regression, with SEs in parentheses.
* Coefficient is twice its SE.
† Coefficient is larger than two and one half times its SE.

year students with college plans or those students who are preparing to take the university examinations.

Investment in shadow education for students is associated with all three measures of the family's SES, even after student and school characteristics are controlled for. Students from wealthier families and families in which the parents have higher levels of education are more likely to purchase shadow education. These effects are moderate, generally adding only 4%–5% to the likelihood of participating in a form of shadow education, but taken together they increase the probability of undertaking shadow education by 12%–15%.[6] For the majority of activities, all three indicators of family SES increase the likelihood of shadow participation.

There are moderate to large gender differences in participation for every activity except after-school classes. Compared with females, males participate more in practice examinations, correspondence courses, and are more willing (54% more) to become ronin after high school. Almost one-half of the male seniors had plans to become ronin if not admitted to university, while only one-tenth of the female students had similar plans.

Scholastic characteristics also influence participation; students with better grades (academic standing) and students in academic curriculum tracks purchase shadow education. The latter effect is large as students in the academic track are 22%–68% more likely than other students to use shadow education, while the largest effect of grades increases participation by only 7%. The curriculum-track effect indicates the strength of the connection between shadow education and educational allocation.

High school reputation has a mixed effect; students from high schools that are successful in university placement use shadow education, except for *juku* (after-school classes). Among these students *juku* is not used because in many prestigious high schools examination preparation activities are now incorporated into an extracurricular schedule, by creating within-school, after-school classes. This shows how pervasive systems of shadow education can in turn influence the instruction and organization of formal schooling.

Finally, living in an urban area moderately increases the use of shadow

---

[6] Effects of the logit coefficients can be represented as an incremental change in the probability of whether a student participates in shadow education as a function of a unit change in a particular independent variable The equation

$$\delta P = \exp(L1)/[1 + \exp(L1) - \exp(L0)]/[1 + \exp(L0)],$$

where $L0$ is the logit before the unit change in x_j, and $L1 = L0 + B_j$ is the logit after the unit change in x_j, estimates this effect on the probability (Petersen 1985).

education, except for the purchase of correspondence courses, the only form of participation not restricted by proximity.

We examine one final issue of participation in shadow education. Some students, who are either not admitted to any university or are admitted but decide to continue to prepare for the examination of a more prestigious university the following year, undertake perhaps the most extensive form of shadow education—the ronin year. Since no limit exists as to the number of times a student can participate in the annual examinations, up to 25% of examination takers in any one year are ronin. While ronin activities can keep the student preparing and competing for the prestigious positions, the financial, social, and emotional costs can be quite high. The average tuition for *yobiko* (private cram school for ronin) ranges from the equivalent of $2,000 to $2,800 per year, but can go as high as $20,000 for *yobiko* designed for entrance to special programs, such as medicine. For some there are additional room and board costs. Other costs for ronin come in the form of emotional stress from constant pressure to prepare, damaged social and peer affiliations, and anxiety about uncertain futures (Tsukada 1988).

To assess which students become ronin, we analyzed those students in the follow-up who, as seniors, had had college plans (69% of the follow-up sample) since this is the pool from which ronin will come. Of this group 27% became ronin the year immediately after high school. Besides the variables used to predict shadow-education participation in high school, we added two additional variables: (1) whether or not the student passed any university entrance examination at the end of high school and (2) a hazard-rate term in the form of a student's probability of being in the follow-up sample.[7] The first measures the opportunity open to students as they decide about the ronin year and the second corrects for any bias occurring from nonrandom sample attrition from the base year to the follow-up year.[8]

---

[7] Since there was attrition between the base year and follow-up samples, we constructed a correction for any selection bias in the data (Berk 1983) A "hazard rate" based on a model of incidental selection (Berk and Ray 1982) was computed and includes indicators of whether the school had large attrition from base year to follow-up, indicators not used in the substantive equations that correlated with family social status, such as whether or not the mother worked, parents' interest in the student's grades, and geographical region. None of the variables used in estimating the hazard rate are used in the substantive equations. The coefficient for this term indicates the amount of change in the dependent variable due to a student's probability of being in the follow-up sample. In two of the three follow-up equations (tables 3 and 4), this proved to be a statistically significant correction for a modest tendency for follow-up students to be more successful in attending university after high school than those base-year students lost through sample attrition.

[8] We drop the student's curriculum track from the remainder of the analysis because, among follow-up students who had college plans, almost all (96%) were in the college-

TABLE 3

INFLUENCES OF STUDENT, FAMILY, AND SCHOOL CHARACTERISTICS ON BECOMING A
RONIN AFTER HIGH SCHOOL AMONG STUDENTS WITH COLLEGE PLANS
(FOLLOW-UP SAMPLE)

Independent Variables	Maximum-Likelihood Parameters	SE
Background factors.		
Gender (1 = male)	1.97†	.20
Academic standing	−.08	.05
Father's education	.10	11
Mother's education	.17*	.08
Family income	.23†	06
High school's reputation	24†	.08
Pass examination (1 = pass)	−3 78†	.20
Rural/urban (1 = urban)	.43*	.19
Hazard rate (follow-up correction)	.98†	27
Intercept	−1.75†	.48
Model χ^2 compared to model with intercept only	796 38	
df	9	
N	1,586	

NOTE.—Maximum-likelihood parameters are derived from logistic regressions
* Coefficient is twice its SE
† Coefficient is larger than two and one half times its SE

As is true of participation in shadow education during the high school
years, family resources have a modest influence on becoming a ronin
after high school (see table 3). Students whose families have more money
and whose parents have more education are more likely to continue in
the most extensive forms of shadow education. These effects are small
however, adding only about 5%–6% to the likelihood that a student
undertakes the ronin year. Family expenditure and student sacrifice for
the ronin year is far more motivated by the gender of the student. Males
dominate the ranks of ronin, and we find that, even after we control for
a number of other factors, when we compare males with females, the
former are 49% more likely to become ronin if this is necessary for contin-
uance in the allocation process.[9] Naturally students who failed examina-
tions during the first year are more likely to become ronin and students

---

bound curriculum during high school. When we add this variable to these models it
does not change the sign or significance of the other effects.

[9] We investigated several family configuration variables such as whether or not the
student was an only child, the oldest child, or had an older sibling of a particular
gender who attended university We found that none of these had any appreciable
influence on the student becoming a ronin.

from rural areas less likely to, because of the higher costs of sending a student to live in the city to attend a *yobiko*. Finally, students from prestigious high schools tend to become ronin, most likely because they are taking the most competitive entrance examinations.

Consequences of Shadow Education

We examine the consequences of participation in shadow education for university attendance at two points in time: (1) the first year out of high school and (2) the second year out of high school. The first point includes all base-year students with university plans and the second point includes only those follow-up-year students with university plans who did not attend university after high school.

Because we have no measure of the quality of the shadow-education experience, our analysis provides a limited assessment of the effect of participation in shadow-education activities on university attendance. We can, nevertheless, determine whether participation in shadow education increases the student's chances of attending university, in addition to including various control variables, such as student, family, and school characteristics, in order to assess the relative contribution of shadow education to allocation outcomes in Japan.

Whether or not a student attends university is regressed twice for each time point. The first logistic equation includes just the block of background variables and the hazard rate for sample attrition, and the second equation adds participation in shadow education. Since various background factors predict participation, we assess the difference between the total effects of background factors and direct effects of background and shadow-education factors to indicate possible indirect effects of the background factors that work through shadow education on university attendance.

Participation in high school shadow education increases the likelihood of university attendance and adds significantly to a model with only background factors ($\delta \chi^2 = 48.45$; $df = 4$; $P < .001$; see table 4). The direct effects of practice examinations and correspondence courses are significant and positive, adding 16% and 25% respectively to the probability of entering a university after high school. Using a tutor decreases one's chances of attendance, which most likely reflects the remedial character of most tutoring compared with the preparatory character of the other forms of shadow education. After-school classes (*juku*) have only a small and nonsignificant effect on attendance, probably because students in better high schools (who tend to do better on examinations) use their high school's after-school program instead of *juku*.

TABLE 4

INFLUENCE OF SHADOW EDUCATION ON UNIVERSITY ENTRANCE FOR STUDENTS WITH
COLLEGE PLANS (Follow-up Sample)

INDEPENDENT VARIABLES	UNIVERSITY AFTER HIGH SCHOOL		UNIVERSITY TWO YEARS AFTER HIGH SCHOOL	
	Total Effect	Direct Effect	Total Effect	Direct Effect
Background factors:				
Gender (1 = male)	2.69†	2.70†	.96†	.16
	(.17)	(18)	(24)	(.30)
Academic standing.	20†	14*	.28†	.27†
	(.05)	(.06)	(.07)	(.07)
Father's education	.19†	15*	02	−.04
	(.07)	(.07)	(.07)	(.08)
Mother's education	.20*	.16	.14	09
	(10)	(10)	(.11)	(12)
Family income	.13†	.12*	.17*	.14*
	(.05)	(.06)	(.07)	(.07)
High school's reputation	.59†	52†	43†	.33†
	(.07)	(.08)	(.10)	(11)
Rural/urban (1 = urban)....... ...	.06	01	39	.28
	(19)	(.20)	(.24)	(.27)
Hazard rate (follow-up correction)	1 28†	1.08†	.43	.16
	(.22)	(24)	(33)	(.38)
Shadow education.				
Practice examination		1.01†		.58
		(22)		(43)
Correspondence course		.64†		−.03
		(.17)		(22)
Private tutor...		− 30		−.79†
		(.25)		(.32)
After-school classes		.30		01
		(17)		(.24)
Ronin...				3.22†
				(57)
Intercept....	−0.37	−1.69†		−1.83*
	(43)	(.49)		(.85)
Model χ² compared to model of				
intercept only	505 99	554 64	80 86	154.11
df.........	8	12	8	13
N	1,135	1,135	511‡	511‡

NOTE —Entries are maximum-likelihood parameters from logistic regression with SEs in parentheses
* Coefficient is twice its SE
† Coefficient is 2 5 times its SE
‡ Students in the follow-up sample with college plans in high school who did not attend a university
immediately after high school.

In the remainder of the direct effects equation, family background, academic ability, and school reputation all influence university entrance. Mother's education clearly works through shadow education indirectly, since adding the shadow-education block drops this coefficient to zero. Males are more likely than females to enter university, probably because of the emphasis placed on males' educational attainment.

The second set of equations in table 4 compares the effects of background factors and shadow education on those students who had plans for college, but did not attend immediately after high school. The dependent variable is university attendance in the second year out of high school. Added to the list of high school shadow education, is whether or not the student had undertaken the ronin year.

Among students, two years out of high school with college plans, high school shadow education had a diminishing effect on attending university, but the effect of being a ronin is dramatic because it increases university attendance by 80%. There is a strong association between the ronin year and continued competition in the race for university; among this part of the follow-up sample 86% became ronin and of these 72% entered college compared with only 6% of the non-ronin from this group. This is particularly true among the males, who are much more likely than females to become ronin. A further 19% of the original ronin continued for a second year to prepare for another round of university examinations.[10] Even though the dominant effect in this model is that of the ronin, shadow education as a group of activities still adds significantly to the background model ($\delta \chi^2 = 73.24$; $df = 5$; $P < .001$).

DISCUSSION

Shadow education in Japan is a pervasive educational activity that primarily focuses on the important transition from high school to university. By successfully negotiating this crucial transition, a student can become a sponsored member of an educated elite. The combination of the potential lifetime pay-offs of this educational transition and the open meritocratic character of the contest greatly increases the potential for shadow education to flourish. As we have seen, many Japanese high school students

---

[10] In analyses not reported here, we have found that shadow education may help students decide which university examinations they should take. From the results of numerous practice examinations and advice of teachers in the *juku*, students can become well-informed about their chances of passing the entrance examination of different universities. Such knowledge is valuable since the scheduling of examinations and the special preparation required limit the number of examinations a student can take.

participate in preparation activities that are not offered in formal schooling, and many high school graduates continue such preparation activities during the two years after high school.

Although rates of participation in various forms of shadow education are uniformly high, there are different patterns of participation. The likelihood of participation increases if the student is male, or has good grades in high school, or is from a higher SES family. These patterns of participation in shadow education reflect broader patterns of stratification by gender, secondary schooling, and SES. Such evidence suggests that participation in shadow education is not a remedial strategy used primarily by students who have difficulty meeting the academic standards of a formal school setting, but is instead a proactive strategy used primarily by students who have already accumulated significant advantages in the formal educational system. Shadow education provides an avenue for parents to enhance their children's chances in the educational allocation contest.

Consequences of participation are modest for some forms of shadow education and large for others. Becoming a ronin is an institutionalized way for one-time losers to stay in the educational allocation contest, and staying in the contest seems to have its pay-offs. Forms of shadow education are designed to help students increase their knowledge in order to improve their performances on the entrance examinations. Gaining admission to a Japanese university is, however, not merely a transfer of family background through investment in shadow education. The meritocratic contest of university admission in Japan emphasizes the academic ability of the student and measures of previous academic performance have large, direct effects on university attendance.

Shadow education is designed to enhance the school careers of students. Its content and existence is tightly coupled to the organization of transitions both within schooling and from school to the workplace. For some societies, the study of shadow education will enhance our understanding of the process by which students are allocated within formal schooling and how social advantages are transferred across generations.

REFERENCES

Amano, I. 1986. "Educational Crisis in Japan." Pp. 23–43 in *Educational Policies in Crisis*, edited by W. K. Cummings, E. R. Beauchamp, S. Ichikawa, V. N. Kobayashi, and M. Ushiogi. New York: Praeger.

———. 1990. *Education and Examination in Modern Japan*. Tokyo. University of Tokyo Press.

Berk, R. 1983. "An Introduction to Sample Selection Bias in Sociological Data." *American Sociological Review* 48.386–98.

Berk, R., and S. Ray. 1982. "Selection Biases in Sociological Data." *Social Science Research* 11:352–98.

Bowles, S., and H. Gintis. 1976. *Schooling in Capitalist America*. New York: Basic.

Bourdieu, P., and J Passeron. 1977. *Reproduction in Education, Society, and Culture*. Beverly Hills, Calif.: Sage.

Brinton, M. 1988. "The Social-Institutional Bases of Gender Stratification: Japan as an Illustrative Case." *American Journal of Sociology* 94.300–334.

———. 1989. "Gender Stratification in Contemporary Urban Japan." *American Sociological Review* 54:549–64.

Cummings, W. 1980. *Education and Equality in Japan*. Princeton, N.J.: Princeton University Press.

Eckstein, M., and H. Noah. 1989. "Forms and Functions of Secondary-School-Leaving Examinations." *Comparative Education Review* 33:295–316.

Franke, W. 1963. *The Reform and Abolition of the Traditional Chinese Examination System*. Cambridge, Mass.: Harvard University Press.

Fujita, H. 1978. "Education and Status Attainment in Modern Japan." Ph.D. dissertation. Stanford University, Department of Sociology.

Katsillis, J., and R. Rubinson. 1990. "Cultural Capital, Student Achievement, and Educational Reproduction in Greece." *American Sociological Review* 55:270–79.

Katsuki, N. 1988. "The Learning Business." *Tokyo Business Today* (January), pp. 34–36.

Kitamura, K. 1986. "The Decline and Reform of Education in Japan: A Comparative Perspective." Pp. 153–70 in *Educational Policies in Crisis*, edited by W. K. Cummings, E. R. Beauchamp, S. Ichikawa, V. N. Kobayashi, and M. Ushiogi. New York: Praeger.

Kondo, S. 1974. "Off We Go to Our Lessons." *Japan Interpreter* 9:15–24

Lin, C. 1983. "The Republic of China (Taiwan)." Pp. 104–35 in *Schooling in East Asia: Forces of Change*, edited by R. M. Thomas and T. W. Postlethwaite. New York: Pergamon.

Meyer, J. 1970. "The Charter Conditions of Diffuse Socialization in Schools." Pp. 564–78 in *Social Processes and Social Structures*, edited by W. R. Scott. New York: Holt, Rinehart & Winston.

———. 1977. "The Effects of Education as an Institution." *American Journal of Sociology* 83:55–77.

Mitchell, R. 1968. *Pupil, Parent and School: A Hong Kong Study*. Taipei: Orient Cultural Service.

Nihon Seishounen Kenkyuujo 1981. *Koukousei Shourai Chousa*. Tokyo: Nihon Seishounen Kenkyuujo.

———. 1984. *Gakkoukyouiku to Sono Kouda*. Tokyo: Nihon Seishounen Kenkyuujo.

Peterson, T. 1985. "A Comment on Presenting Results from Logit and Probit Models." *American Sociological Review* 50:130–31.

Rohlen, T. 1983. *Japan's High Schools*. Berkeley and Los Angeles: University of California Press.

Rosenbaum, J., and T. Kariya. 1989. "From High School to Work: Market and Institutional Mechanisms in Japan." *American Journal of Sociology* 94:1334–65.

Shimahara, N. 1984. "Socialization for College Entrance Examinations in Japan." *Comparative Education* 20:253–66.

Sweeting, A. 1983. "Hong Kong." Pp. 272–97 in *Schooling in East Asia: Forces of Change*, edited by R. M. Thomas and T. W. Postlethwaite. New York: Pergamon.

Tsukada, M. 1986. "A Factual Overview of Japanese and American Education." Pp. 96–116 in *Educational Policies in Crisis*, edited by W. K. Cummings, E. R. Beauchamp, S. Ichikawa, V. N. Kobayashi, and M. Ushiogi. New York: Praeger.

———. 1988 "The *Yobiko:* The Institutionalized Supplementary Educational Insti-

tution in Japan: A Study of the Social Stratification Process." Ph D. dissertation. University of Hawaii, Department of Sociology.

Turner, R. 1960 "Sponsored and Contest Mobility in the School System." *American Sociological Review* 25:855–67.

U.S. Department of Education. 1987. *Japanese Education Today*. Washington, D.C.. Government Printing Office.

Ushiogi, M. 1986. "Transition from School to Work: The Japanese Case." Pp. 197–209 in *Educational Policies in Crisis*, edited by W. K. Cummings, E. R Beauchamp, S. Ichikawa, V. N. Kobayashi, and M. Ushiogi. New York: Praeger.

White, M. 1987. *The Japanese Educational Challenge*. New York: Free Press.

SYMPOSIUM ON SCALING OCCUPATIONS

Scaling the Intergenerational Continuity of Occupation: Is Occupational Inheritance Ascriptive after All?[1]

Steve Rytina
McGill University

A new scaling for detailed occupational categories is prepared by extracting a symmetric scaling of intergenerational continuity (SSIC) from a 308 × 308 mobility table derived from the NORC General Social Survey. The resulting SSIC scores are based solely on observed mobility among detailed occupational categories. These scores are quite similar to socioeconomic index (SEI) scores and somewhat less similar to NORC occupational prestige scores. The SSIC values closely match average offspring's SEIs and average father's SEIs for each category, although no information about SEI or its components is employed in their calculation. This fundamentalist approach to measuring inheritance without a priori scaling reveals that mediation of occupational inheritance by education is outweighed by direct "ascriptive" transmission.

I. INTRODUCTION

The intergenerational continuity of occupation is the most extensively investigated question in American sociology. Yet investigations require severe abstraction from the underlying phenomena. The raw machine-readable data are typically classifications into the detailed occupational categories of the U.S. Census Bureau. But the raw codes must be transformed before analysis. In one tradition, the occupations are assigned to

ranks most commonly based on Duncan's (1961) socioeconomic index (SEI) or upon prestige scores (Siegel 1971). In the other, the detailed occupations are grouped into roughly ranked functional categories broadly derivative of Edwards (1938).

This article describes an alternative to such sacrifice of detail. I use a very large (and very sparse) contingency table based on 308 detailed occupational categories to produce a ranking of occupation that optimally captures intergenerational continuity. I call these optimal ranks symmetric scaling of intergenerational continuity (SSIC) scores. While the ranks are derived from the details of a table, the scale values serve as a preliminary to regression analysis of status reproduction. Thus the proposed approach can be viewed as a partial synthesis of the two conventional alternatives.

The SSIC is a distinction that leads to a large difference: it fails to sustain the central finding of status transmission research; that is, that education predominates as an intervening variable mediating occupational rank in successive generations, so that individual "achievement" prevails over "ascription" or direct, unmediated inheritance. But in the metric of SSIC, education's role as mediator is sharply diminished. The greater continuity of occupation described by SSIC is dominated by ascription.

This sharp contrast with prior findings also applies to Hout's (1988) contingency-table finding that "expanding universalism" prevails among the well educated. Yet my analysis is based on the same data as those used by Hout. (The contrasting findings are briefly explored below.) Thus I will show that this central finding of educational mediation is scale dependent, and I will present a first look at a revised picture of intergenerational occupational continuity.

The substantive stakes are high. And a critical issue is the warrant for replacing established conventions with an unfamiliar and abstract procedure. I will argue for two complementary advantages, one analytic, the other empirical.

First, SSIC scores are more fundamental. While accepted occupational scales made regression analysis of mobility possible, they enter as more or less arbitrary conventions. In contrast, SSIC operationalizes a central assumption; that is, that nominal occupational categories are arranged into a hierarchy that constrains intergenerational movement. The analytic notion that movement among categories constitutes a change in rank suffices to define relative ranks for categories. Therefore, one can eliminate the imposition of an a priori metric from the study of mobility.

Second, SSIC results can be viewed, to a good empirical approximation, as correcting the widely employed SEI scaling for modest but significant defects in representing intergenerational continuity as a linear

1659

regression. Thus SSIC differs from SEI in pattern and degree where occupations are linked by ascent and descent to more or less favorable origins and outcomes than SEI rank would suggest. Ironically, the defects in particular SEI values follow from imposing SEI as a measuring rod for mobility.

The irony reflects the tension between the two rationales. If the analytic warrant is accepted, then empirical comparisons with previous conventions are irrelevant. But I will adopt a skeptical attitude and emphasize the empirical. The strong conventions governing occupational scaling are the foundation of much work, and modifications should face a severe test. The powerful rationale for procedures that are scale blind is counterbalanced by the opacity of abstract algorithms. Accordingly, I will place much emphasis on the close similarity of SSIC results with established conventions and show that the new scaling can be seen as "corrected values" that correspond with an empiricist critique of SEI. In short, the analytic warrant will be underplayed until the empirical case is laid out.

II. OCCUPATION, OCCUPATIONAL SCALING, AND INTERGENERATIONAL CONTINUITY

In the real world, people have jobs that lead to differential access to valued resources and privileges (cf. Jencks, Perman, and Rainwater 1988). But social researchers have long drawn on the tradition of occupational coding (Conk 1978) to distinguish among classes of positions in the division of paid labor. The raw material is rendered machine readable when respondents' answers to questions about their jobs are transformed by coding teams into the detailed occupational classification (or DOC) of the U.S. Census Bureau (for reliability of this procedure, see Bielby, Hauser, and Featherman [1977]). There are more than 400 such categories, but the distribution is highly uneven and the largest categories contain the bulk of the cases.

These categories are inescapably ad hoc. At least roughly, they differentiate function in the division of paid labor. They reflect task description (What kind of work do you usually do?) and organizational context (e.g., manufacturing vs. nonmanufacturing is frequently contrasted). They do not directly reflect authority, power, social class, or any of the more popular generative sources of a unitary social hierarchy. And yet various aspects of hierarchy are clearly implicated. Professions requiring extensive training, skill, or both are differentiated from one another and from jobs less demanding of human capital. Different occupational skills (butcher, baker, and candlestick maker, except that the last are not evident and the butchers are divided into manufacturing and nonmanufac-

turing) are distinguished, and these distinctions often delimit differential access to social advantage.

In general, the information implicit in the detailed categories is not analyzed owing to the limitations of established methodology. Instead, the categories are grouped by two broad methods into equivalence classes, and the distinguishing information is discarded.

Grouping finer distinctions into broader functional categories is an extrapolation of the initial theme of similarity in job tasks. Such coarse schemes are the usual preliminary to contingency table analysis. The simplest such schemes distinguish farm, blue-collar, and white-collar occupations. The more elaborate versions, such as the 17-category scheme of Blau and Duncan (1967), make further distinctions of function and skill level. These distinctions roughly correspond to modal educational qualifications and economic rewards. But the internal differentiation of occupations and persons that is thereby ignored is enormous.

The resulting contingency tables are now most often analyzed by log-linear and related methods. The earliest analytic methods did not incorporate any scaling information. Latter-day association models (Goodman 1979; Hout 1983) arrange the nominal categories into numerically differentiated orders. Hout (1984) has accomplished a partial merger of distinct traditions by integrating external scales (e.g., status as measured by average SEI) with association models.

Log-linear analysis and related methods are widely thought vulnerable to empty cells in contingency tables.[2] This has often led researchers to aggregate categories. The resulting combinations (into occupational groups) of combinations (of detailed occupational categories) of combinations (of occupational titles) of jobs are highly abstract representations of occupational differentiation. Somewhat greater detail is incorporated in linear, additive styles of analysis, but the additional discriminating power comes at some cost: a price is paid in the form of stronger assumptions made prior to empirical analysis.

One of the great wellsprings for research on occupational continuity was the development of "continuous" scales for the detailed occupational categories. Duncan's (1961) SEI was based on data on income and education levels from the 1950 census. He derived a scale that closely approximated survey measures of subjective occupational standing. One great

[2] For an illustrative prescription see Wasserman and Faust (1989, p 21). This would seem to indicate that the maximum-likelihood proposals of Gilula and Haberman (1986) and related proposals by Goodman (1986, 1987) are not applicable. This is somewhat true, but there is a strong and useful convergence between some of the present techniques and some of the proposals in the literature, which I intend to spell out in later work.

advantage of his scale was that it could be extended to many more occupational titles. Siegel (1971) ingeniously merged subjective results from many studies to achieve comparably broad coverage for what are commonly called "prestige scores." There have been further developments, notably by Treiman (1977), Nam and Powers (1983), and Stevens and Featherman (1981) (along with independent efforts for Great Britain), but the most widely employed are closely related to the original proposals.

As Duncan noted back in 1961, the scales that result are not continuous but impose equivalence classes on the detailed occupational categories (which are themselves equivalence classes imposed on the even more differentiated occupational titles). For example, packers and wrappers, sewers and stitchers, and guards and watchmen are equated in the SEI rank of 18.[3] Right next door, in rank 19, are bartenders, not-specified operatives, miscellaneous operatives, machine operators miscellaneous specified, and carpenters. These diverse possibilities are equated (and nearly equated) for the purpose of analysis.

Intergenerational occupational continuity is not directly assessed by any standard model. Strictly speaking, Blau and Duncan (1967) and their intellectual descendants have only measured the degree of continuity of occupational SEI. Proponents of prestige scales within the same tradition have only examined the continuity of subjective evaluations of job titles. It is therefore improper to equate scale scores with "occupation." Correlations of scale scores are one step removed from the continuity of "occupation" in all of the differentiated detail that the standard data archives would allow.

But these methods were, of course, a great advance and a practical necessity. The blooming buzzing confusion of detailed functional differentiation carries within it almost no directly accessible continuity. For example, in the 308 category occupational mobility table to be examined below only 501 of 7,965 or 6.3% of the offspring occupy the same detailed category as their male parent.

The assumption, by now amply documented by empirical research, is that status continuity consists in the fact that offspring are concentrated in occupations "close" to that of their male parent. But "close" requires some metric of social distance. The scales in popular use are the tools that allow us to "see" a move from stitcher to watchman as status continuity (in SEI = 18) and from stitcher to bartender (from SEI = 18 to SEI = 19) as almost no change at all. The degree of continuity re-

[3] The SEI and prestige values for this analysis were taken from App. P of the codebook for the OCGII (Data Program and Library Service 1983). To minimize clumsy wording, the third digits (beyond the decimal place) of scale values are ignored in discussion, although the full values were used in calculations

vealed by correlations of rankings support the assumption of constraining distance. (For a detailed discussion defending this now-taken-for-granted assumption see Blau and Duncan [1967, chap. 4].)

The several accepted ranking schemes differ in detail, and they reveal different degrees of status continuity. These differences are generally left unexplored, much less explained. Systematic empirical comparisons are reported in Featherman, Jones, and Hauser (1975), Featherman and Hauser (1976), and Stevens and Featherman (1981). The prevalent interpretation is that prestige is more error ridden. But Stevens and Featherman illustrate the equally prevalent ambivalence. "We do not attach any substantive difference to these differences [between attainment equations employing prestige and various SEI measures]. They are impossible to sort out without convincing evidence that what one index measures is conceptually and operationally distinct from another" (1981, p. 377). Since differences are slight, and not systematically understood, there exists a strong pressure to employ standard rankings in empirical research. An implicit taboo against idiosyncratic revisions minimizes an unmanageable source of noncomparability across studies.

This strong conventional strand is both convenient and empirically grounded. Differences in scales lead to modest differences in results, and there exist no obvious grounds for identifying observed differences with theoretically meaningful distinctions. After all, the SEI was constructed to replicate prestige judgments; the scales are intentionally similar. Insofar as there is no dissent against the original constructions (although some proponents of SEI oppose the use of prestige for attainment studies), they serve as conventions. The cumbersome lists of scale assignments can be machine coded and, afterward, taken for granted. Accordingly, there is scant evidence of concern for the fine details of occupational placement, in large part because there is no obvious criterion for criticism.

III. THE DATA

Since one aim of the present analysis was comparison with Hout (1988), I started with the same subset of the NORC General Social Survey (GSS; see Davis and Smith 1986) that he did. These data consist of 8,623 cases of persons in the labor force, age 25–64, surveyed from 1972 to 1986, who were assigned SEIs for both father and self.

Since the intermediate object of this analysis is the detailed occupation, the data were restricted to 7,998 respondents for which both reported occupation and father's occupation had at least one observation, not normally from the same respondent.[4] Seven of the remaining 315 occupa-

----

[4] Ten cases were eliminated for missing values on education.

tions were at the outer limit of just one father and one offspring. I further eliminated seven sparsely populated occupations that resulted in "misleading" outliers (see App. A), resulting in 7,965 cases and 308 occupations. Over 92% of the potential case base was retained.

The resulting mobility table is sparse, to put it mildly. There are 94,864 cells of which only 4,672 are nonzero. The mean value for the nonempty cells is only 1.71; there are but .084 cases per cell for the full table. Accordingly, I did not further exclude observations on racial or gender grounds.[5]

IV. DETECTING DEFECTS IN THE DETAILS

Measurement cannot be properly isolated from application within a pattern of relationship. Proponents of status attainment have endorsed SEI as the superior measuring rod for assessing intergenerational change or mobility. The generic or central application of SEI is representation of inheritance as a statistical regression. Taking this as the core provides a basis for assessing validity of specific scale values.

An implicit assumption is that DOCs that share SEI levels will not differ in inheritance outcome, except for sampling error. The conditional means, or average rank associated by inheritance with detailed occupations, should be a linear function of SEI value, specifically the predicted values from regression.

But this is a testable hypothesis. Each father's DOC has an empirical average for offspring SEI, and that average can be compared with the predicted value based on father's SEI. The corresponding F-test is equivalent to testing the incremental impact of adding dummy variables for each DOC to the SEI regression. For the data in the 308 × 308 table employed in this analysis, the F-test for no difference for offspring SEI among fathers' DOCs is 1.372 (with 307 and 7,690 df; $P < 2 \times E^{-5}$). Departures from SEI linearity far exceed expectations based on sampling error.

From the point of view of a tabular accounting, this is only half the pattern, namely, the half corresponding to outflow. There is a second regression that "postdicts" father's SEI from offspring's, and this regression corresponds to inflow. While the first regression describes descent, the second describes ascent. Of course, the latter is statistically dubious on well-known grounds. But it is nevertheless indicative of nonlinearities consequent on SEI scaling. The F-statistic for no difference in average

---

[5] To avoid unduly complicating the parallel algebra and numerical checks used to gain confidence in these procedures, I also did not apply weights for the oversampling of blacks for 1982, which insured that my matrix had integer entries.

origin among offspring DOCs net of SEI is 1.545 ($P < 2 \times E^{-9}$). Thus ascent shows even more marked departures from linearity than descent.

The very tiny P values mask the fact that occupations vary widely in the degree to which SEI provides an adequate summary of descent and ascent. To bring this out, I note first that each F is (approximately) the mean of 308 squares of frequency standardized deviations from linear prediction associated with each detailed occupation. Furthermore, the deviations for ascent and descent are largely independent for each occupation, since the ascendants and descendants are different cases except for the 6.3% of the cases that fall on the diagonal. Therefore, one can associate, with each occupation, an approximate z-statistic that is based on the sum of the deviation from linear prediction for ascent and descent respectively.[6]

Examination of these (almost) z-statistics indicate that 261/308 of the occupations have associated (two-tailed) attained levels of significance greater than .05. However, the 29 occupations for which .05 > P > .01, the 11 occupations for which .01 > P > .001, and the 7 occupations for which P < .001 are clear evidence that offspring from a substantial fraction of DOCs systematically do better or worse in outcome and origin than can be implied by the linear SEI characterization.

A further subtlety follows from the implications of relative frequency. Among the 80 occupations with 50 or more representatives in the sample, one-quarter, or 20, lead to P < .05, which far exceeds chance expectation. The remaining 27 instances of P < .05 observed among the 228 less frequent occupations also exceed the chance expectation of 11.4, but the excess is less dramatic.

Thus most occupations have patterns of ascent and descent adequately summarized by SEI. Fully three-quarters of the 80 for which there are sufficient data for a reasonably powerful test yield results consistent with the null hypothesis. However, there is also evidence that systematic deviations exist; this is most striking for frequent occupations, but is by no means absent among those with few representatives in the sample.

The same pattern can be found in metric differences. The overall position of a DOC with respect to inheritance can be summarized by the average deviation about the respective (fathers' and offspring's) means of the SEI of ascendants and descendants of a DOC. This will be discussed below as the "rough-and-ready" scale.

For each DOC, the regression frame establishes an expectation: the frequency-weighted average based on linear prediction and postdiction. This expectation is, of course, a function of the SEI for each DOC. The "metric discrepancy" between observed and expected values reflects the

[6] The sum of the z's is divided by the square root of 2 to standardize to unit variance.

consistency of SEI rank and empirical inheritance. Such differences should be compared to the standard deviation (SD) of predicted values of 7.72 SEI scale units. Forty-three occupations have a difference of more than 10 SEI scale units. Although none of these is among the 80 occupations with 50 or more representatives (hereafter referred to as the 80 "largest" occupations), 16 of the 43 combine sufficient frequency and deviation to reject the null at the .05 level. Furthermore, 60 other occupations have a difference of 5–10 SEI scale units. Twenty-two of these differences are statistically significant, including all of the 11 from the 80 largest occupations with discrepancies in this range. The median discrepancy is 3.85. Thus many of the discrepancies are quite small. But there are some occupations more than one SD out of alignment, and many of these contain enough observations to render such differences unlikely by chance.

This evidence is fuzzy, but telling. Average origins and outcomes are, for most occupations, accurately summarized by SEI. Yet, for a notable minority, most visible among occupations with high sample frequency, there are departures. Some are no doubt merely sampling fluctuations. But it is terribly unlikely that all are. Thus one must conclude that not all SEI values accurately reflect the empirical pattern of inheritance. In an added irony, which also magnifies the obscuring fuzz, it is SEI values that provide the measuring rod for origins and destinations that partially invalidates some of their kindred.

In sum, a closer look at the details reveals differences among DOCs not accounted for by SEI. In terms of ascent and descent, some do better and some do worse than their SEI rank would imply. Identity in SEI value does not delimit equality of access to and from other occupational ranks. This heterogeneity within ranks of SEI is information about occupational continuity that is sacrificed by the use of SEI scaling.

V. EXTRACTING A SCALE OF INTERGENERATIONAL CONTINUITY

The ideal would be to adjust scores using the details of the social mobility table. A means for doing this is provided by the technique of canonical scaling that is described by Kendall and Stuart (1979). (It was first proposed by Hirschfield [1935] and applied by Fisher [1940] and Maung [1940]). These methods have been applied to mobility tables based on coarse occupational classifications by Klatzky and Hodge (1971), Hope (1972), Duncan-Jones (1972), Featherman et al. (1975), and Sawinski and Domanski (1987). The results for U.S. data have converged in finding that the emergent order for the broad categories closely parallels average SEI rankings for the occupations grouped into the categories. Feather-

man et al. further showed that the parallel was fuzzier for occupational prestige, and they concluded that prestige scores "are fallible estimates of the socioeconomic statuses of occupation" (1975, p. 329).

Canonical scaling is described in the sources given above. It requires the calculation of eigenvalues and eigenvectors of a transformed variant of the mobility table. The eigenvectors corresponding to the largest (first) eigenvalue are standardized scores, or weights, for father's occupation and offspring's occupation that maximize the correlation between the two (accordingly, the scales are labeled EIG1f and EIG1o).[7] Since no other assignment of numbers to father and offspring DOCs can produce a higher correlation, any particular assignment, including prestige or SEI, must result in a lower correlation. Therefore the algorithm generates an upper bound on the father-offspring correlation that could be generated from the data in the given table.

One possible drawback of this approach is that the scale scores of fathers' DOCs and offspring's DOCs need not be equal for asymmetric tables (although they have previously been reported as very similar for tables of aggregated occupational categories). In this instance the father and offspring scales only correlate .565.[8] The notion that DOCs are located in a hierarchy that constrains mobility entails a symmetric scale that applies the same rank to both fathers and offspring.

Values for such a scale were obtained by a second algorithm (described in App. B) that directly maximized the father-offspring correlation subject to the constraint of equality between father and offspring scalings. This variant on canonical scaling produces a similar result. The scale that emerges correlates .961 with the frequency-weighted average of the asymmetric variants. For 188 of the 308 occupations, the symmetric scale value lies between the asymmetric fathers' and offspring's values. The symmetric scale results in a father-offspring correlation of .450, which is only slightly less than the canonical correlation of .489.[9] I call the numeri-

---

[7] There is a first dimension that records category frequency, so that the first substantively interesting dimension under consideration here is technically the second. Also see App. A.

[8] The correlation based on offspring frequencies is .565. Unless noted, this base will be used since the data is most nearly a random sample for offspring. One can also use occupation (r = .412), or father frequencies (r = 819), as a unit of analysis. An indication of the source of this pattern, which is recurrent, is that r = .772 for the 80 largest, with occupation as the unit of analysis. Smaller occupations have a larger number of wilder results; the huge number of farmers among fathers leads to high correlations.

[9] I carried out these analyses using various software. App. B indicates how SPSS and its case-sequential kindred can be used. Most of the development was carried out in the Gauss language (Edlefsen and Jones 1984; details and programs are available from the author on request).

cal placement of DOCs along this scale, which maximizes the father-offspring correlation, the symmetric scaling of intergenerational continuity. I refer to the values as SSIC scores. Strictly speaking, they are SSIC scores only for the data on which they are based, and other more or less similar arrays would arise from different samples. They array detailed occupations as nearer and farther from each other along a dimension that corresponds to maximum stickiness or constraint in observed movement.

The only information that enters the analysis is the purely nominal father-offspring co-occupancy of position in the division of paid labor. The result has a distinctive conceptual purity. It directly operationalizes the hypothesis that movement among positions in the division of labor is constrained by a hierarchy. This widely shared assumption can be used to produce a pattern that is conceptually independent of the causes and consequences of such a hierarchy.

The resulting map reveals the pattern of neighborhoods in the division of paid labor to which offspring are probabilistically confined by the "accident of birth." Locations that are "near" and "far" are inferred from frequency of movement. Distance and movement are duals. The result, from one angle, is simply the most parsimonious (one-dimensional) description of the constraint governing intergenerational stability. But conceptually, this description can be taken as *the* primary pattern of ascription governing occupational placement, insofar as "ascription" means assignment on the basis of origin conceptually isolated from other attributes of the individual. The counterpart of this conceptual purity is the elimination of possible contamination that might enter if anticipated causes and consequences of occupational ascription were forced in as a priori indicators of occupational rank.

But conceptual purity carries with it considerable risk of pragmatic nonsense. The (rhetorical) table underlying the analysis below is extremely sparse. The algorithms are merely numeric, and therefore stupid. Conceptual independence of other notions (and metrics) of hierarchy might lead to gross inconsistency with the useful information that I deliberately did not incorporate. The resulting description could easily enough be a mere curiosity (or even monstrosity) unrelated to more pragmatic efforts to impose some pattern on the complexity of the division of paid labor. In short, it remains to be seen whether conceptual virtue will trash common sense.

VI. EMPIRICAL VALIDATION

Validity may be pursued at two levels. At a global level, the correlation of new results with received scales can be taken as validating the latter. This interpretation would parallel that of previous reports (Featherman

TABLE 1

CORRELATIONS AMONG VARIOUS RANKINGS OF DETAILED OCCUPATIONAL CATEGORIES

	Siegel	SEI	Rough and Ready	SSIC	EIG1*o*	EIG1*f*
Siegel	1.0	.867	.765	.702	.655	.507
SEI	.806	1.0	.854	.819	.757	.572
Rough and ready	.700	.891	1.0	.951	906	601
SSIC	592	839	.957	1 0	.942	.661
EIG1*o*	.528	.774	883	.928	1 0	565
EIG1*f*	.565	800	888	.924	.806	1.0

NOTE.—This table is based on the 7,965 cases in the 308 × 308 mobility table used to derive SSIC scores Results above the diagonal are based on offspring frequencies, results below the diagonal are based on father frequencies. SEI refers to socioeconomic index scores Rough and ready refers to the mean deviation of SEI for fathers and for offspring. EIG1*f* and EIG1*o* are the first dimension canonical scorings for fathers and for offspring.

et al. 1975; Klatzky and Hodge 1971) based on similar methods applied to aggregated tables which, of necessity, reported correlations of canonical scores with average SEI for the aggregates. But my analysis also provides access to the lower level details of scale values. Insofar as the rerankings correspond with demonstrable defects in accepted values, then the new scale may lay claim to a superior validity.

Table 1 reports correlations between various rankings of the 308 detailed occupational categories. The SSIC scores extracted from mobility data alone correlate .819 with SEI. The correlation with prestige (Siegel scores) is a more modest .702.[10] This means that something closely akin to SEI is implicit in the most parsimonious summary of the disaggregated details of the full mobility matrix. To a considerable degree, the pattern of movement among nominal occupational categories is one of distance along SEI. The less robust performance of Siegel scores replicates the parallel finding of Featherman et al. (1975).

One could simply take this as further validation of SEI. But a more radical stance is available. SEI, after all, is not constructed to capture or reflect the detailed pattern of mobility. It is an exterior construct adopted for mobility studies as a convenient approximation to occupational rank.

On the other hand, SSIC emerges from the empirical pattern of mobility, subjected to the hypothesis that the pattern is one of distance along

---

[10] The results in this section are generally stronger for the larger occupations. For example, for the 80 most frequently observed occupations, each of which occurs in 50 or more observations, the correlation of SSIC with SEI is .890.

a dimension. This operation reveals a hierarchy of occupations that is quite similar to hierarchies derived from quite different, and empirically independent, criteria. But SEI is a key to bringing order to a potentially very messy domain. Before an alternative is advocated, it is useful to address suspicions that an arcane computer routine has yielded mysterious numbers not readily assigned meaning within the established frame of reference. Table 2 contains some raw materials for "making sense" of the modifications of rank entailed by the new procedure.

In Table 2 I report correlations of various scales with several more accessible empirical intermediaries. Every occupation can be ranked by the average SEI for the fathers of its incumbents. The correlation of this with other scales is reported in the first row, labeled PA AVE. Similarly, every occupation can be ranked by the parallel average SEI for offspring who share a common occupational origin (labeled OFF AVE).

The results show that both of the scalings derived from the mobility table are more closely correlated with the details of the continuity of SEI than is SEI. Specifically, SSIC correlates .906 with average offspring's SEI, while SEI only correlates .844. Also, SSIC correlates .923 with average fathers' SEI, where SEI correlates .811. Similar advantages over SEI are revealed for corresponding asymmetric scales (although not for unlike combinations). The differences are modest. But it is remarkable that the derived scales, which come from procedures blind to SEI values, more closely reflect the pattern of SEI inheritance than does SEI.

The combined average of all SEIs linked to a given occupation by both ascent and descent (with SEIs taken as deviations about the offspring's or fathers' mean as appropriate) was contrasted in Section IV with regression predicted values, and there the difference was called the "metric discrepancy." The averages of linked SEIs are a scaling of occupation in terms of SEI continuity. It appears in both tables 1 and 2, with the label "rough and ready."

TABLE 2

CORRELATIONS OF VARIOUS RANKINGS OF DETAILED OCCUPATIONAL CATEGORIES
WITH AVERAGE OFFSPRING'S SEI AND AVERAGE FATHERS' SEI

	Siegel	SEI	Rough and Ready	SSIC	EIG1f	EIG1o
PA AVE	.715	811	.953	.923	.531	939
OFF AVE	.639	.844	928	906	.925	.779

NOTE —This table is based on the 7,965 cases in the 308 × 308 mobility table used to derive SSIC scores. PA AVE is the average fathers' SEI for members of each occupation, while OFF AVE is the average offspring's SEI for fathers in each occupation Calculations for PA AVE are based on offspring marginal frequencies, while those for OFF AVE are based on father marginal frequencies

As described in Section IV, the rough and ready scale is largely parallel to SEI but with some significant differences. The values can be seen as corrections to SEI incorporating the details of SEI continuity. The corrections are not massive, insofar as rough and ready correlates .854 with SEI. But it is more closely related to SSIC with a correlation of .951. Thus rough and ready serves as an intuitively accessible bridge between SEI and the modifications that reflect the detailed patterns of intergenerational continuity.

Of particular interest are the differences of SEI and SSIC. These differences are often small, but they correlate .861 with the "metric discrepancy" between observed (rough and ready) and predicted SEI, in which predicted SEI is a linear function of SEI. To a very good approximation, SSIC modifies SEI values only where SEI poorly reflects continuity of SEI across generations. The material in Section IV showed that these discrepancies cannot be laid to sampling error. Thus SSIC can be viewed as "correcting" SEI by repositioning occupations to a closer correspondence with the details of ascent and descent.

It must be stressed that this result emerges even though the SSIC procedure was quite blind to SEI or any other exterior ranking information. Indeed, the result is somewhat ironic. The SSIC could be motivated as using the details of the sparse mobility table to derive the most parsimonious or unidimensional description of flows among detailed occupations. Yet the result more nearly corresponds with the detailed pattern of ascent and descent for SEI than does SEI. The evidence that reveals this equally suggests that SEI placement is slightly misleading for capturing the pattern of occupational mobility.

The irony is that this evidence against SEI follows from adopting it as a provisional criterion for rank, so that one may obtain average origins and outcomes. Thus the basis for claiming superior empirical validity for SSIC is the conjunction of SEI, a linear inheritance representation, and the distinction among detailed occupations. A proponent of unmodified SEI would face the dilemma that the standard linear inheritance model is misspecified with respect to details. But SSIC leads to new placements of occupation in close correspondence with the SEI-based indications that occupations are misplaced.

The contrast for mobility studies may be worth serious attention. Occupations that share SEI rank but ascend to and descend from different levels are patterns of movement that are more apparent than real. Insofar as these results are averages, they are, by standard convention, the "structural" expectations for individuals, that are logically prior to further individual characteristics. Indeed, differences in average outcome for those descending from distinct origins is the underlying core operational meaning of the central notion of intergenerational ascription. My

evidence shows that SSIC is superior to SEI as a summary of the relative positions of DOCs in this regard.

VII. THE INTERGENERATIONAL CONTINUITY AND TRANSMISSION OF OCCUPATIONAL RANK

Table 3 displays the amount of intergenerational continuity revealed through the various alternative codings of occupation. The general pattern is quite striking. The sequence from Siegel to SEI to rough and ready to SSIC to canonical scores (labeled EIG1) is one of increasing continuity. From small to large, Siegel results in a correlation ($r = .256$) that is only 76% as large as that for SEI ($r = .338$). Rough and ready leads to an increment over SEI of 26% ($r = .426$), while SSIC results in an increment over SEI of 33% ($r = .450$). The increment for SSIC over Siegel is 76%. The most extreme result, of course, is for the asymmetric canonical scorings. This correlation of .489 is the greatest that can be produced by any scoring. It is 45% greater than that for SEI and 91% greater than that for Siegel.

If one prefers the standard of explained variance, the contrast is necessarily greater. Explained variance for SSIC is 309% of that for Siegel and 177% of that for SEI. The canonical scaling registers explained variance of 365% of that for Siegel and 209% for that of SEI. By any conventional criteria, these results show that traditional methods understate the continuity of occupation to a considerable degree.

For about a generation, the measurement of occupational standing has not been an end in itself but a means to the study of intergenerational transmission. The general finding is well known: occupational rank is mainly transmitted through intervening attainment of education. In statistical terms, the indirect effect through education accounts for the greater part of the observed intergenerational correlation of occupational rank. While the meaning of this result is contested, there appears to be no remaining dispute about the empirical situation.

TABLE 3

CORRELATIONS MEASURING THE INTERGENERATIONAL CONTINUITY OF
OCCUPATIONAL RANK

	Siegel	SEI	Rough and Ready	SSIC	EIG1
OFF rank	.256	338	.426	.450	.489

NOTE —This table is based on the 7,965 cases in the 308 × 308 mobility table used to derive SSIC scores Correlation involving Siegel does not include 177 cases with fathers or offspring who are current or retired members of the armed forces

TABLE 4

CORRELATIONS OF DIFFERENT MEASURES OF OCCUPATIONAL RANK
WITH RESPONDENT EDUCATION

	Siegel	SEI	Rough and Ready	SSIC	EIG1
PA rank	313	.401	422	.423	.417
OFF rank	.585	.587	.567	.550	.521

NOTE —This table is based on the 7,965 cases in the 308 × 308 mobility table used to derive SSIC scores. Correlations involving Siegel do not include 177 cases with fathers or offspring who are current or retired members of the armed forces.

As a first step to reexamining this result, table 4 describes the correlation of fathers' and offspring's rank on various metrics with education. Once again, the hierarchy of metrics shows a pattern. "Better" measures of occupational continuity produce stronger relations between father's rank and offspring's educational attainments. Siegel (1971) produces a correlation of only .313. The use of SEI produces a substantially higher .401. Rough and ready, at .422, nearly ties SSIC, at .423, and the canonical scoring (labeled EIG1 in the table) trails off slightly to .417.

The correlation of education with offspring rank shows a reverse pattern. Siegel and SEI are highest and nearly identical at .585 and .587, respectively. At .567, rough and ready is somewhat smaller, while SSIC yields .550, and the canonical scoring is the smallest, although still quite large, at .521.

Figure 1 and table 5 display the implications for the status transmission process. Passage up the hierarchy of scales results in a consistent and quite marked enhancement of the effect of father's rank, or "ascription." This is accompanied by a weaker but still striking decrease in the net impact of education and a consequent decrease in the indirect effect, or "achievement." The magnitude of the direct effect more than doubles. The predominance of "achievement" over "ascription" that is apparent with the traditional measures is replaced by a predominance of "ascription" over "achievement." Revised scaling turns this familiar pattern on its head.[11]

---

[11] As a reviewer commented, verbalizing this contrast is somewhat a matter of "half empty" vs. "half full." Though the numbers are as reluctant as ever to speak for themselves, most readers will have no trouble comprehending such familiar diagrams and assigning their own summary language. And, as can be plainly seen, education by no means drops out of the picture, but is demoted from the main attraction to an important sideshow.
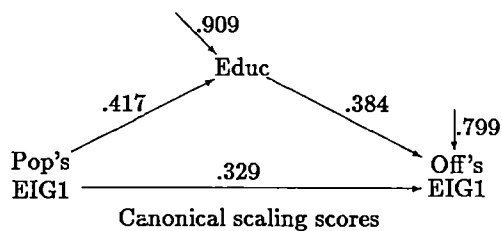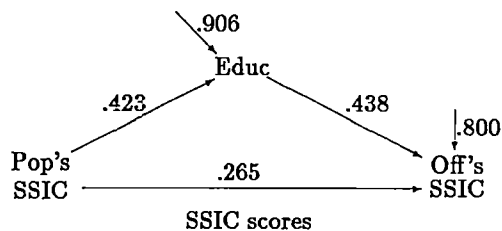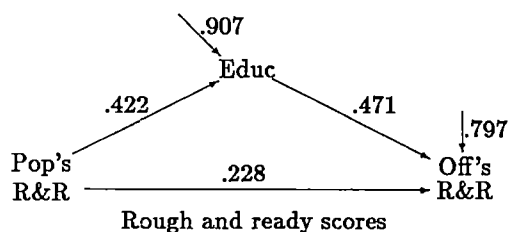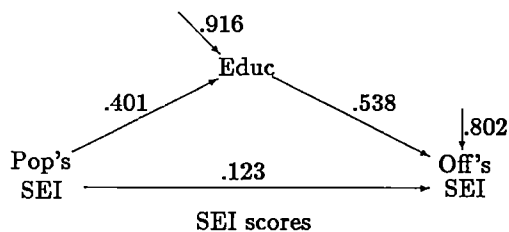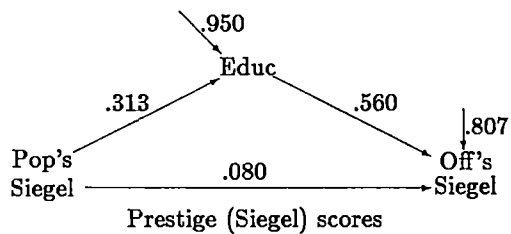
FIG. 1.—Path diagrams of the role of education in intergenerational continuity of occupational rank. Pop's = fathers'; Off's = offspring's; R&R = rough and ready. Based on the 7,965 cases in the 308 × 308 mobility table used to derive SSIC scores. See table 1 for meaning of variable labels. Results for Siegel do not include 177 cases with fathers or offspring who are current or retired members of the armed forces.

TABLE 5

TOTAL OCCUPATIONAL CONTINUITY AND NET DIRECT EFFECT OF FATHER'S RANK ON
OFFSPRING'S RANK FOR DIFFERENT MEASURES OF RANK

Scale	Continuity Correlation	Percent Direct Effect
Siegel	256	31.4
SEI	.338	36 3
Rough and ready	.426	53 4
SSIC	.450	58.8
EIG1	.489	67 3

NOTE.—This table is based on the 7,965 cases in the 308 × 308 mobility table used to derive SSIC scores. Result for Siegel does not include 177 cases with fathers or offspring who are current or retired members of the armed forces.

Thus the traditional measures of occupational rank understate occupational inheritance in several complementary ways. First, because they do not directly reflect the pattern of movement among occupations, they understate the "nearness" of fathers' and offspring's detailed occupations. Siegel does an especially poor job, while SEI comes fairly close to the optimum. Second, the correlation of education with occupational rank declines.

These might readily have been expected. The new measures are based on optimizations, so the increase in inheritance follows by construction. And one might expect the idealism implicit in prestige judgments to inflate the apparent contribution of such a morally attractive basis for social rank as education, while SEI directly incorporates education (albeit of occupational peers) as a component. Indeed, the small size of the change indicates that such contamination as there might be is modest.

The more surprising result is that father's rank is more closely correlated with offspring's education for the scales that more accurately capture occupational continuity. This result is fully contingent and not due to construction. Substantively, this confirms that offspring education is an asset that is acquired differentially by those from occupations lower or higher along the axis of continuity. It is consistent with the interpretation that high-rank fathers possess some generalized means to improve offspring's stratificational chances. More subtly, this inferred capacity to facilitate accumulation of human capital is (slightly) better captured by SSIC than by traditional measures.

A concrete example will help indicate how dramatic this difference is. The four-year contrast of college versus high school graduation is 1.305 SDs. In the SSIC model, this translates into a net gain of .5715 occupation standard units. To achieve this same gain by "virtue" of family

background requires 2.156 standard units of father's rank. Thus a physician's offspring (origin $z = 2.65$) with a high school degree has about the same occupational expectations as the offspring of locomotive engineers, roofers ($z = .48$), and banktellers ($z = .49$) who finish college.

In the SEI model, college versus high school graduation translates into .7015 occupational units. But the difference in social origin necessary to obtain the same advantage would be 132.3 SEI scale units. In the SEI image, college education would more than bridge the 93 point gap between the most and least advantaged occupations, while in the SSIC image it would overcome less than half the rungs in the occupational ladder.

The new picture that emerges is less morally appealing than the previous versions. Offspring of occupations higher in the "contest" for status maintenance acquire more education. At the same time, rank is retained in considerable measure by factors independent of education, "because" adult rank is less strongly correlated with education. The differences are modest, but are large enough that the noneducational path of the continuity of social rank is allowed to emerge.

VIII. ARE THE RESULTS ROBUST?

The falsification of a well-established finding in Section VII is necessarily tentative. The thesis of education as mediator is based on many studies and several samples. My demonstration is based on only a single sample. Replication on larger samples and comparison of results across different samples will be required to clarify the extent to which the finding is local to the GSS.

But two possible defects can be dealt with immediately. First, the use of a sample to fit scales raises the possibility that SSIC values are contaminated by "overfitting" accidental fluctuations in the raw mobility patterns under analysis. Second, the results might be somehow inflated by failing to restrict the sample to nonrural occupations.

Three pieces of evidence can be adduced against the thesis of overfitting. First, as reported above, changes in rank correspond rather closely with demonstrable defects in SEI as a characterization of typical origin and destination. These defects occur across occupation sample frequencies. There are no good grounds for dismissing the defects, or the parallel shifts in rank, particularly for the larger occupations. And it is these that largely determine global results like correlations.

Second, if overfitting were occurring due to sampling fluctuation, one would expect that the less frequently observed occupations would show

more wobble. But the correlation of observation frequency and the absolute difference in scale values between SEI and SSIC is $-.126$. Thus less frequent occupations tend to be placed slightly closer to their SEI position.

Third, the analysis reported here is only the final result of my expanding slowly to include rare occupations. Any distortion due to infrequent categories should have been less when these were excluded, but no such pattern occurs. As an illustration, within the 308 × 308 table is a 91 × 91 subtable for the 91 occupations that each have (in the subtable) at least six fathers and at least six offspring and at least 25 observations for fathers and offspring combined. For this subset, the SSIC correlation increases slightly, from .45048 to .45106, which is the opposite change than would be expected from the overfitting thesis. Indeed, the correlation for SEI is less stable, with the subsample yielding .35585 compared to .33822 /in the full sample. For the summary of percent direct effect, net of education, the SSIC results again show a slight increase in the subsample, to .60537 compared with .58833, while SEI is again less stable, going from .36256 overall to .39260 for the subsample.

The issue of nonrural occupations might be raised because the rural occupations are often excluded in a tradition that dates back to Blau and Duncan (1967), and they are outliers for both SEI and SSIC. Exclusion decreases the SSIC correlation by 5.7% to .42419, but decreases the SEI correlation by 8.6% to .30925. The percent direct effect for SSIC is slightly higher, at .59987, while that for SEI is slightly lower, at .35145. The changes are minor and clearly indicate that the results do not turn on the treatment of rural occupations. While I cannot now address every imaginable possibility, the brief summary of these three pieces of evidence does indicate that the results of this analysis are at least not due to certain "obvious" sources of possible distortion.

IX. ARE THE NEW RESULTS SOUND EVIDENCE OF GREATER ASCRIPTION?

To this point, the primary grounds offered to justify SSIC have been empirical. A skeptical attitude toward a new approach is justified (or even dictated) by the costs implicit in rejecting well-established, widely employed conventions. But with evidence in hand that differences in scale values correspond to a pattern of empirical flaws, and that a major substantive claim is at stake, I will now attempt the stronger argument that SSIC is also conceptually superior, particularly with respect to the achievement/ascription contrast.

The first issue is the relative status of the different scales. Prestige, and especially SEI, are the standards on which established findings are based. Each rests, to some degree, on an a priori rationale. It is true that each requires qualification, as in, "Ascription with respect to scale Y," but each rests on criteria independent of the data on hand.

But close examination of the underlying rationales reveals a central difficulty. Prestige, honorably enough, is solidly grounded in the Weberian corpus. It operationalizes the analytic assertion of a pattern of social distance among occupations, and this pattern rests on a consensus about honor. Average perceptions of social honor (and more arguably, average answers to the actual question[s] about social standing; cf. Hope [1982]) correspond to the intrinsic logic of Weber's position. [12] But measured prestige falls short of SEI in revealing pattern, and this comparative weakness in "explained variance" has lead to disapproval (e.g., Stevens and Featherman 1981) and disuse. Unfortunately, SEI's "construction does not appear to be grounded in sociological theory" (Hodge 1981, p. 414). There are powerful pragmatic grounds for preferring SEI, but no compelling theoretical rationale exists. Bluntly stated, it is ad hoc.

However, SSIC is on a theoretical footing more comparable to that of prestige, and is more empirically powerful than SEI. The empirical power, of course, is insured by construction and the magnitude of the difference has been explored above. It is the theoretical rationale that now deserves attention.

The metaphor of stratification is intrinsically geometric. The notion presupposes that the division of labor arrays (labor-selling) persons along a hierarchy. One conceptual property is that the stratification of families (restricted by the usual fiat to fathers) confines descendants to occupations more or less adjacent to their originating rank. The counterpart of probabilistic confinement to layers is movement across distances, but motion in a social hierarchy presupposes a metric.

Occupational mobility, in turn, presupposes a metric over occupations. The intrinsic metric is that which describes the pattern of constraint, or the neighborhoods of occupations to which origin probabilistically confines offspring. More simply, ease and difficulty of intergenerational transition is distance made manifest. On this preconception, the pattern of mobility is the empirical realization of a hierarchical arrangement of occupations.

The SSIC is the constructive realization of this analytic or preempirical conception. It operationalizes the hypothesis of constraining distance by

[12] "Analytic" is used to designate the preempirical cutting up of the world on the basis of pure conceptual distinction. In these terms, ideal operational measures are those that correspond most closely with the conceptual notions that motivate them.

generating a (best one-dimensional) map from the flows among nominal categories. Just as averaged sentiments derive meaning from the analytic assumption of consensus about honor, the simple map of flows is the empirical counterpart of the metaphor of movement among layers.

"Ascription," in turn, may be analytically defined as the pattern of assignment in the division of paid labor consequent on the position of family of origin. As a summary of the pattern of adjacency or restricted outcomes governing intergenerational reproduction of the division of labor, SSIC directly operationalizes ascription. Therefore, SSIC is not merely a possible metric for assessing occupational achievement versus ascription; it is the ideal metric.

Some may sense a danger of tautology in the application of this exhaustive account to the standard statistical rendition of this issue. Insofar as ascription is further identified with the direct effect of father's rank, and achievement with the mediating effect of education, it might seem that optimizing constraint with respect to father's rank unfairly tips the balance against education, and for ascription.

But there is no tautology involved. The formula of achievement versus ascription is improperly suggestive of a horse race, where each element should be given an equal opportunity for success. More properly, the issue is whether factors interpretable as reflecting individual accomplishments, like educational attainment, account in the sense of path analysis for the baseline correlation indexing intergenerational ascription. There is nothing "fair" about deflating the baseline to reduce the constraint that factors of achievement must explain.[13]

The greatest advantage of SSIC is that it empirically exhausts the analytic notion that outcome is constrained by origin. It summarizes the empirical pattern by which this occurs. It therefore provides a full accounting of occupational ascription, which is the appropriate explanandum for any attempt to represent inheritance as a consequence of individual attainment.

One can make no comparable claims for SEI. First, it has no conceptual resonance with ascription, but is ad hoc. Second, it has empirical defects such that occupations sharing SEI levels give rise to different average outcomes. Such differences are "ascribed," that is, are assign-

---

[13] Some reviewer comments seemed to reflect the belief that if one is to "rescale" occupation, one should do the same for education. As the text indicates, such symmetry is misplaced. Occupation must be scaled, while education does not require such treatment. In any case, it makes little difference. To allow education as much slack as possible, I fitted it with 19 dummies for the 20 distinct levels. The increment to R^2 was just over 1%, indicating that education is almost exactly linear in its effect. The residual direct effect of SSIC was diminished from .265 to .260, which is negligible.

ments to locales of the division of labor that are consequent on the accidents of birth. Recruitment shows a parallel deflation of real versus "predicted" fluidity. On this view, the SEI and education regression is, at best, an imperfect approximation employing a somewhat arbitrary baseline.

Comparison with conventional scales leads to a decisive conclusion. "Mediation by education" and "occupational hierarchy as constraint on mobility" are empirically incompatible. One can maintain the thesis of mediation by education only by imposing metrics that do rather poorly at representing constraint on mobility. Insofar as the latter is most closely parallel to the analytic definition of ascription, the present evidence is that occupational ascription cannot be explained as a consequence of occupational assignment based on educational attainment.

X. A PROGRESSIVE INTERPRETATION OF SUCCESSIVE OCCUPATIONAL SCALES

Scaling does not occur in a vacuum, but is purposive. One formulation of the driving interest behind scaling is that occupations are loci of assets and advantages. This formula suggests that they both offer attractions for which people compete and resources that underlie the mechanism by which rank is reproduced.

From this viewpoint, prestige provides a means to assess ascription based on an implicit claim of a Weberian view: status groups act collectively to conserve honorific standing. The relatively poor correlation of prestige and SSIC can be viewed as refuting this claim. The pattern of occupational inheritance is not one of conserving honor.

But if this thesis fails, why should SEI reveal more pattern? After all, SEI was devised as a convenient approximation to prestige judgments. And one would expect an alternative calibrated to an underlying dimension to result in attenuated correlations.

An alternative interpretation is causal: SEI consists of two components reflecting the income and education levels of occupational peers. One can construe this as a combination of monetary and nonmonetary rewards. Thus SEI is then a measure of the advantages for which individuals compete. But the causal interpretation is that SEI measures sources of power—both the power that stems from market position and the power that stems from human capital assets. These assets translate into causal control or influence over the outcomes of offspring.

On this view, SEI was an inspired guess that came closer to revealing the hard facts driving rank reproduction than did prestige. It was a better measure because the assets through which honor is sustained are more potent than the respect that tends to accompany them. Thus SEI pene-

trated closer to the causal essence by more nearly reflecting the sources of differential power that determine differential outcomes.

Rough and ready takes us one step further. Some occupations are differentially successful in recruiting from and sending to loci of advantage/disadvantage. Employing this index takes us one step further from limiting assumptions about the sources of power, and one step closer to the pattern of outcomes that are the product of differential ability to affect the fate of offspring.

The SSIC is a step further still. If one regards SEI as a good first approximation to the means and ends of status competition, then rough and ready is a corrected approximation that averages over errors in SEI, and SSIC is the continuation of that averaging process.[14]

Implicitly, SSIC will reflect additional dimensions of occupational assets beyond those incorporated in SEI. It responds to the general hypothesis that differences in outcome have systematic causes among such mental, moral, and material concomitants of occupation as allow incumbents to variously cripple, enhance, or otherwise constrain access to outcomes by offspring. While SEI implicitly limits contrasts to two (very central) causal dimensions, SSIC is sensitive to such differences as emerge from examination of the detailed averages.

This interpretation suggests that SSIC is the appropriate framework for uncovering the interpretable differences through which rank of family of origin is a cause of outcome. And it suggests that new issues and new answers, possibly consistent with the power imagery suggested here, will have to emerge. For SSIC demonstrates that the most widely accepted locus of continuity—education—is not the predominant generative source for the intergenerational stability of occupational rank.

XI. DOES A COLLEGE DEGREE BREAK THE ASCRIPTIVE LINK?

In a recent attempt to update the Blau and Duncan (1967) interpretation, Hout (1988) used modern tabular techniques and recent samples to advance a strong but modified version of the achievement or "expanding universalism" thesis. According to Hout, "The results here support a different view, one that sees universalism within those positions that have been most selective in drawing human resources from colleges and universities but particularism within those positions that have traditionally selected from among those workers who lack advanced credentials" (1988, p. 1381).

---

[14] The SSIC algorithm terminates when the "metric discrepancy" is zero. Similarly, the first step, when SEI is the start value, is essentially equal to "rough and ready," while subsequent steps respond to such discrepancies as still remain.

Hout's interpretation has many facets. It is based on an extended series of applications of his SAT (status, autonomy, training) model to different data sets and to subsamples in those data sets. His work represents the state of the art in contingency-table methods. But the keystone of his analysis is that the impact of origin status on destination status is "zero for college graduates (at least among male college graduates over 35 years old)" (p. 1384). The qualifying phrase is absent from the conclusion, where Hout says, "the effect of origins on destinations differs by level of education. The extreme case is college graduates. For them, occupation status is independent of origin status. This finding provides a new answer to the old question about education's overcoming disadvantaged origins. A college degree can do it" (p. 1391). It is this summary statement, which I think contains Hout's central claim, that I will examine here.

Although Hout used tabular methods, he directly compares his estimates of effects to those found by manipulating covariances (i.e., regression and correlation). He draws on the same data employed in my analysis (although of course he did not restrict his sample to occupations represented by both fathers and offspring). The first aspect of his key claim is that the effect of origins on destinations differs by level of education. Table 6 displays this pattern.

TABLE 6

THE CORRELATION OF OFFSPRING'S OCCUPATIONAL RANK WITH FATHER'S
OCCUPATIONAL RANK FOR DIFFERENT LEVELS OF EDUCATION

Schooling Completed	Scaling Method	r	S/E	t	P	N of cases
0–7 years	SEI	.1366	.053	2.55	0107	342
	SSIC	.3316	.051	6 50	<.0001	
8 years	SEI	.0808	.055	1.48	.1387	333
	SSIC	.2877	052	5 48	< 0001	
9–11 years	SEI	.1349	.030	4 46	< 0001	1,073
	SSIC	.2729	029	9 29	<.0001	
High school	SEI	1760	019	9.38	<.0001	2,750
	SSIC	2947	018	16 17	< 0001	
13–15 years	SEI	.1493	.023	6 08	<.0001	1,620
	SSIC	3032	024	12.80	< 0001	
College grads	SEI	0687	.023	2 95	.0031	1,847
	SSIC	.2844	.022	12 74	<.0001	

NOTE.—This table is based on the 7,965 cases in the 308 × 308 mobility table used to derive SSIC scores.

1682

In table 6 I report the correlation of origin and destination calculated within the several strata of education. The data manipulations for table 6 differ somewhat from Hout's. Hout recoded occupational titles into broad categories and then assigned 14 different levels of aggregate SEI, while my table 6 is based on the full disaggregated ranks. Where possible, he distinguished genders and races. He restricted analysis to the 1972–85 samples (except for a few tabulations), while table 6 includes the 1986 data. But the results are similar and show "universalism" for college graduates. The effect of origins on destinations approaches zero for these folk. The actual value of the correlation (.0687) is significantly different from zero by statistical criteria, but it records an association that is negligible by any other standard.

Table 6 also contains the comparable result for SSIC. The difference is dramatic. The effect of origins among college graduates is far from negligible. And, in contrast with SEI, SSIC effects vary little across educational ranks. This means that the ordinary path model, which assumes no interaction, is an adequate description. Indeed, the SSIC results are nearly the same as in the path model except that a slight magnification of the residual direct effects results from the washing out of effects of education differences within educational categories.

The difference between analyses is dramatic: the keystone of the interpretation stands or falls on choice of scale. Hout's analysis is an elegant example from a well-developed empirical tradition. But the time may be ripe to reevaluate the foundations and conventions on which that tradition rests.

XI. CONCLUSION

The confirmatory character of much of the preceding deserves emphasis. First, while the details of SEI inheritance depart somewhat from a strict linear representation, differences were restricted to a minority of occupational categories. Second, the axis of continuity revealed by analysis of the 308 category mobility table was relatively close to SEI and further from prestige, as suggested by previous studies of aggregated tables. Third, SSIC was shown to reflect the corrections to SEI implied by examination of SEI continuity for detailed occupational distinctions. This is quite remarkable in light of the fact that SSIC scores are calculated without any information about SEI or its components.

These results largely confirm that SEI is an excellent a priori approximation to the locus of immobility among occupations. At the same time, they indicated room for improvement. The rough and ready scores were one step, but these averages of SEI origins and destinations were also a bridge to the full-fledged maximization of SSIC.

However, the confirmatory character of my investigation stopped abruptly at the mechanism of intergenerational transmission. The widely analyzed data used here shows the expected major role of education as rank transmission mechanism. But recasting rank onto the fundamental axis of intergenerational continuity revealed that direct transmission, unmediated by education, plays an even larger role.

The replacement of accepted conventions on which this rests should not be undertaken lightly. The ritual call for further research has more than the usual urgency. But SSIC is in tune with important deeper themes. It directly operationalizes the shared metaphor of mobility as the counterpart of a hierarchy that defines near and far across occupations. It can be seen as modest corrections that supplant the necessity of a priori assumptions by drawing on the full information present in the averages of flows among detailed occupations. The result is a cleaner and more direct reflection of ascription than was heretofore available.

A new, somewhat unfamiliar picture is apparent in the previously discarded details. The nominal coding of detailed occupational categories reveals considerable constraint on mobility. This constraint is substantially stronger than was apparent with conventional ranking schemes. Offspring of the most favored occupations are the most successful in obtaining education, and such success matters. But the continuation of occupation across generations is substantially independent of the educational accomplishments of offspring.

APPENDIX A

On the Deletion of Some Outlying Occupations

The largest table examined for this project was one of 99,225 cells and only 7,998 observations, based on the 315 occupations meeting the inclusion criteria. Some might think it somewhat foolish to blindly apply numerical algorithms to such a sparse structure. There turns out to be solid ground for this hunch.

In actual execution, the analyses reported here were slowly built up beginning with the 36 largest occupations, and gradually extending to those with minimum marginals of 6, 5, . . . , 1. The results were very stable, but anomalies eventually emerged.

To overcome these, seven occupations, involving 33 cases, were ejected for producing wild values and, not incidentally, impeding convergence of the various algorithms (the details will be supplied by author on request). Briefly, rare occupations that are predominant sources (or sinks) for other rare occupations can cause the algorithms to oscillate and can result in extreme values, contrasting a few occupations (or one occupation) with the rest as an undifferentiated lump.

Accordingly, the omitted occupations show up as, for example, the extreme anchors of vectors corresponding to three additional larger eigenvalues when the full 315 occupations are subjected to canonical analysis. Such results are sensible after a fashion; an occupation that is purely self-reproducing would require an extreme value on its own dimension. But various near approximations to this involving rare occupations do not merit much substantive weight.

The largest occupation excluded was the nine observations of funeral directors, involving eight fathers and three offspring, of whom two were following their fathers' line of undertaking. The most extreme value on SSIC occurred for the four glaziers, consisting of three fathers and two offspring, with one inheritor. Ultimately, such near approximations to self-reproducing occupational communities are realistically regarded as occupying their own distinctive axis of mobility removed from the rest of the sample. But the sample is too thin to justify pursuing the matter.

I chose to omit "ill-behaved" occupations rather than clutter the report with "silly" dimensions, even though the results make sense on close examination. Somewhat similarly, results concerning secondary dimensions that seem of no substantive interest have been omitted in the final revisions.

There are a variety of ways to describe how one arrives at the "real" dimensions. They are local solutions in the neighborhood of standard scales. They are the dimensions that emerge when analysis is restricted to the more frequent occupations. They are the dimensions that correspond to the largest canonical correlation associated with a reasonably even dispersion of scale values across the entire set of occupations. And they are strict maxima for the set of 308 occupations employed in the analysis.

These extreme results justify the suspicion that mechanical application of the present procedures could lead to nonsense. Yet it does not require very deep knowledge of the occupational structure to pick out the meaningful results. The "lesser" dimensions reported in this article are prominent features of the data, but without prior substantive commitments and a certain measure of patience one might not detect them.

APPENDIX B

An Algorithm for Maximizing the Correlation by Application of Equal Scores to the Rows and Columns of a Contingency Table

In the following, I will use {rule for all i} to denote a vector and {rule for all i, j} to denote a matrix. Thus P = {sample proportion of offspring in occupation j with fathers in occupation i i, j = 1 ... N} denotes a mobility table. Its ith row is P_t = {P_{ij} j = 1 ... N}, and jth column is

$P'_j = \{P_{ij}, i = 1 \ldots N\}$. Let $P_o = \{P_{\cdot i}, i = 1 \ldots N\}$ and $P_f = \{P_{i\cdot}, i = 1 \ldots N\}$ denote the vectors of marginal proportions of offspring and fathers.

The problem is to find a vector $\tilde{w}$ such the correlation r of father with offspring is a maximum. Let $\tilde{w}_o = \{w_i - P'_o w, i = 1 \ldots N\}$ stand for the vector of deviations around the offspring mean, with elements $[\tilde{w}_o]_i$, $\tilde{w}_f$ for the deviations around the fathers mean, and $V_o(w)$ and $V_f(w)$ for the variances based on offspring and father frequencies. Then r is $\tilde{w}'_f P \tilde{w}_o / \sqrt{V_o(w)\,V_f(w)}$.

Rather than directly maximizing r, it is easier to work with

$$\log(r) = \log(\tilde{w}'_f P \tilde{w}_o) - (1/2)\{\log[V_o(w)] + \log[V_f(w)]\}.$$

A typical approach would be to find the gradient and the hessian so as to apply a Newton-Raphson procedure. Unfortunately, the hessian is analytically singular, so that the necessary inverse cannot be obtained. I therefore took an alternative approach.

Assuming that one is in the neighborhood of a solution, the object is to find $\bar{\epsilon}$ such that $w_{t+1} = w_t + \bar{\epsilon}$ is closer to the maximum. The derivative with respect to $\bar{\epsilon}$ is

$$
\begin{aligned}
\frac{\delta \log(r)}{\delta \bar{\epsilon}} = &\ \{(\tilde{w}'_o P_{\cdot i} + \tilde{w}'_p P'_i)/\tilde{w}'_f P \tilde{w}_o,\ i = 1 \ldots N\} \\
&- \{P_{\cdot i}[\tilde{w}_o]_i/V_o(w),\ i = 1 \ldots N\} \\
&- \{P_{i\cdot}[\tilde{w}_f]_i/V_f(w),\ i = 1 \ldots N\} \\
&+ \{\epsilon_i[2(P_{ii} - P_{\cdot i}P_{i\cdot})]/\tilde{w}'_f P \tilde{w}_o,\ i = 1 \ldots N\} \\
&- \{\epsilon_i P_{\cdot i}(1 - P_{\cdot i})/V_o(w),\ i = 1 \ldots N\} \\
&- \{\epsilon_i P_{i\cdot}(1 - P_{i\cdot})/V_f(w),\ i = 1 \ldots N\}.
\end{aligned}
$$

The first three terms are strictly a function of w, while the last three include ϵ. Setting the derivative equal to zero and solving for $\bar{\epsilon}$ yields the steps for the algorithm. Thus

$$\epsilon_i = \frac{(\tilde{w}'_o P_i + \tilde{w}'_f P'_i)/\tilde{w}'_f P \tilde{w}_o + P_{\cdot i}[\tilde{w}_o]_i/V_o(w) + P_{i\cdot}[\tilde{w}_f]_i/V_f(w)}{P_{\cdot i}(1 - P_{\cdot i})/V_o(w) + P_{i\cdot}(1 - P_{i\cdot})/V_f(w) - 2(P_{ii} - P_{\cdot i}P_{i\cdot})/\tilde{w}'_f P \tilde{w}_o},$$

describes the amount to be added to update the scale value for the ith category.

Although the expression may appear formidable, it is easily implemented in case-sequential software such as SPSS (1985). The key is to associate with each occurrence of an occupation, either for father or offspring, such quantities as the sample frequencies for father and for offspring, and scale dependent quantities, such as $\tilde{w}'_f P'_i$. The latter is the

sum of all father scale scores for incumbents of occupation i divided by the sample size. One can form such sums in an auxillary file and then associate them with father and offspring instances by matching. Once the pieces are assembled, the update for each occupation can be calculated for every case in which it occurs.

A similar trick can be used to extract the principal eigenvectors. These are the limit of taking the preceding sum, converting it to an average over the fathers associated with a given occupation by ascent, and then taking the average for fathers of the resulting scale score of offspring. It can be shown that this "forward and back" averaging is equivalent to repeated multiplication by the matrix whose eigenvectors yield the canonical scores, and thus it is equivalent to a well-known technique for extracting the principal eigenvector.

REFERENCES

Bielby, William, Robert Hauser, and David Featherman. 1977. "Response Errors of Black and Nonblack Males in Models of the Intergenerational Transmission of Socioeconomic Status." *American Journal of Sociology* 82 (May): 1242–88.

Blau, Peter M., and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York: Academic.

Conk, Margo. 1978 "Occupational Classification in the United States: 1870–1940." *Journal of Interdisciplinary History* 9:111–30

Data Program and Library Service. 1983. *Occupational Changes in a Generation, 1973: A User's Guide to the Machine Readable Data File* Madison, Wis.. Data Program and Library Service.

Davis, James A., and Tom Smith 1986. *General Social Survey, 1972–86: Cumulative Data* (MRDF). NORC ed. Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.

Duncan, O. D. 1961 "A Socioeconomic Index for All Occupations." Pp. 109–38 in *Occupations and Social Status*, edited by A. Reiss New York: Free Press

Duncan-Jones, P. 1972. "Social Mobility, Canonical Scoring, and Occupational Classification." Pp. 191–210 in *The Analysis of Social Mobility· Methods and Approaches*, edited by Keith Hope. New York· Oxford University Press.

Edlefsen, Lee E., and Samuel D. Jones. 1984. *Reference Guide to Gauss* Kent, Wash.: Applied Technical Systems.

Edwards, Alba M. 1938. *A Social-Economic Grouping of the Gainful Workers of the United States*. Washington, D.C. Government Printing Office.

Featherman, David, and Robert Hauser 1976. "Prestige or Socioeconomic Scales in the Study of Occupational Achievement?" *Sociological Methods and Research* 4 (May): 403–23

Featherman, David, F. Lancaster Jones, and Robert Hauser. 1975. "Assumptions of Social Mobility Research in the U.S.: The Case of Occupational Status." *Social Science Research* 4:329–60.

Fisher, R. A. 1940. "The Precision of Discriminant Functions." *Annals of Eugenics* 10:422–29.

Gilula, Z., and S. Haberman. 1986 "Canonical Analysis of Contingency Tables by Maximum Likelihood " *Journal of the American Statistical Association* 81·780–88.

Goodman, Leo 1979. "Multiplicative Models for the Analysis of Occupational Mobil-

ity Tables and Other Kinds of Cross-classification Tables." *American Journal of Sociology* 84:804–19.

———. 1986. "Some Useful Extensions of the Usual Correspondence Analysis Approach and the Usual Log-linear Models Approach in the Analysis of Contingency Tables." *International Statistical Review* 54:243–70.

———. 1987. "New Methods for Analyzing the Intrinsic Character of Qualitative Variables using Cross-Classified Data." *American Journal of Sociology* 93 (November): 529–83.

Hirschfield, H. O. 1935. "A Connection between Correlation and Contingency." *Proceedings of the Cambridge Philosophical Society* 31:520–24.

Hodge, Robert M. 1981. "The Measurement of Occupational Status." *Social Science Research* 10:396–415.

Hope, Keith. 1972. "Quantifying Constraints on Social Mobility: The Latent Hierarchies of a Contingency Table." Pp. 121–90 in *The Analysis of Social Mobility: Methods and Approaches,* edited by Keith Hope. New York: Oxford University Press.

———. 1982. "A Liberal Theory of Prestige." *American Journal of Sociology* 87 (March): 1011–31.

Hout, Michael 1983. *Mobility Tables.* Beverly Hills, Calif.: Sage.

——— 1984. "Status, Autonomy, and Training in Occupational Mobility." *American Journal of Sociology* 89 (May): 1379–1409.

———. 1988. "More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s." *American Journal of Sociology* 93 (6): 1358–1400.

Jencks, Christopher, Lauri Perman, and Lee Rainwater. 1988. "What Is a Good Job? A New Measure of Labor-Market Success." *American Journal of Sociology* 93 (May): 1322–57.

Kendall, Maurice, and Alan Stuart. 1979. *The Advanced Theory of Statistics,* vol. 2. New York: Macmillan.

Klatzky, Sheila, and Robert M. Hodge. 1971. "A Canonical Correlation Analysis of Occupational Mobility." *Journal of the American Statistical Association* 66:16–22.

Maung, K. 1941. "Measurement of Association in a Contingency Table with Special Reference to the Pigmentation in Hair and Eye Colours of Scottish School Children." *Annals of Eugenics* 11:189–223.

Nam, Charles B., and Mary G. Powers. 1983. *The Socioeconomic Approach to Status Measurement.* Houston: Cap and Gown Press.

Sawinski, Zbigniew, and Henryk Domanski. 1987. "Dimensions of Social Stratification: The Empirical Invariance." *European Sociological Review* 3·39–53.

Siegel, Paul. 1971. "Prestige in the American Occupational Structure." Ph.D. dissertation. University of Chicago, Department of Sociology.

SPSS. 1985. *SPSSX User's Guide,* 2d ed. New York: McGraw-Hill.

Stevens, Gillian, and David Featherman. 1981. "A Revised Socio-economic Index of Occupational Status." *Social Science Research* 10 364–95.

Treiman, Donald J. 1977. *Occupational Prestige in Comparative Perspective.* New York: Academic.

Wasserman, Stanley, and Katherine Faust. 1989. "Canonical Analysis of the Composition and Structure of Social Networks." *Sociological Methodology* 19:1–42.

How Not to Measure Intergenerational Occupational Persistence[1]

Robert M. Hauser and John Allen Logan
University of Wisconsin—Madison

This article is an invited critique of Rytina's proposal that canonically weighted scales for detailed occupations be used to assess intergenerational occupational persistence. These canonical scales yield higher intergenerational correlations than the Duncan SEI primarily because they capitalize on chance fluctuations in the data used to construct them. Thus, there is no reliable basis for revising previous findings about the relative importance of ascription and achievement in the stratification process. Canonical scaling (and related log-linear and log-multiplicative models) are no more than a useful adjunct to models that locate the sources of intergenerational occupational persistence in the measurable characteristics of occupations and persons.

In "Scaling the Intergenerational Continuity of Occupation," Steve Rytina (1992, in this vol.) attempts to improve the measurement of occupational social standing and to revise prior findings about the persistence of occupational standing across generations by specifying the scale of occupational standing with a canonical weighting of detailed occupational categories. As he notes, Klatzky and Hodge (1971) carried out a canonical analysis of Blau and Duncan's (1967) 17 × 17 classification of

occupational mobility among U.S. men in 1962;[2] they noted the similarity between socioeconomic characteristics of occupations and their estimates of canonical scores for the 17 occupational categories. Rytina confirms this key finding of Klatzky and Hodge by analyzing mobility among 308 detailed occupational categories in the cumulative General Social Survey (GSS), 1972–86. He also proposes a new scale of occupational status, the symmetric scaling of intergenerational continuity (SSIC), which is based on constrained canonical weights for detailed occupations.

Rytina's analysis yields three striking empirical findings. First, the observed correlation between offspring's and father's status increases from .256 using the NORC (Siegel 1971) prestige scores and .338 using the SEI (Duncan 1961) scale to .450 with the author's preferred SSIC scale, in which the canonical weights (scale values of occupations) are constrained to be equal in the generations of parents and offspring. Rytina says that the SSIC corrects the SEI and, we assume, prestige scores, by improving the correspondence of occupational ranks with "the details of ascent and descent." Thus, he finds that "continuity of occupation" is understated when traditional scales are used.

Second, aside from the larger estimate of the intergenerational correlation, Rytina finds that educational attainment is much less important in mediating the persistence of occupational social standing from one generation to the next. This reduction occurs through an increasing dependence of schooling on father's occupational standing and a roughly compensating decrease in the influence of schooling on offspring's occupational standing, along with the larger overall correlation between father's and offspring's occupational standing. That is, the indirect effect of father's occupational standing is roughly constant, but the direct effect increases. This, he says, contradicts widely accepted findings by showing that ascription is more important than achievement in explaining intergenerational occupational persistence.

Third, contrary to Hout's (1988) analyses of (nearly) the same data, Rytina finds that the dependence of offspring's occupational social standing on parental occupational standing is maintained among college graduates. Hout (1988, p. 1391) had argued that, for college graduates, "current occupational status is independent of origin status. This finding provides a new answer to the old question about education's overcoming disadvantaged origins. A college degree can do it." Rytina finds that origins have almost no effect on destinations for college graduates when

[2] Rytina does not mention two applications of canonical analysis to the estimation of trend in intergenerational occupational mobility (Hauser and Featherman 1977, pp. 169–87; Featherman and Hauser 1978, pp. 63–138)

occupations are indexed by the SEI, but that there is a strong effect if occupations are indexed instead by the SSIC. The correlation between father's occupational standing and offspring's occupational standing among college graduates is only .0687 when the SEI is used, but it is .2844 when the SSIC is used.

Finally, and somewhat independently of his empirical findings, Rytina argues for the theoretical superiority of his canonical scale in assessments of the intergenerational "continuity of occupation." He argues that the SSIC is ideal for assessing whether ascription or achievement is more important.

In our discussion, we use "occupational standing" as a generic term covering prestige scales, socioeconomic indexes, and Rytina's canonical scores for occupations. We agree with Jencks's (1990) emphasis on the substantive difference between occupational prestige and occupational socioeconomic status and, thus, on the different interpretations of their intergenerational correlations. Occupational prestige scores measure prestige, and their validity does not stand or fall on the size of intergenerational prestige correlations. However, in this discussion, we ignore differences between prestige scales and socioeconomic indexes or canonical scores and focus on those between socioeconomic indexes and canonical scores. Also, we have restricted our discussion to intergenerational correlations of occupational characteristics and ignored correlations of other variables. We think it entirely possible that intergenerational correlations of other social and economic characteristics or combinations of such characteristics, including occupations at different stages of the life course, may be higher or lower than the correlations of characteristics of occupations considered by Rytina. Finally, other measures of the influence of family origins on social and economic success, like the similarity of siblings, may suggest greater or lesser levels of intergenerational persistence. In short, we do not want our discussion of intergenerational occupational correlations to be read as more general remarks about the process of social stratification.

WHAT'S WRONG WITH THE SSIC?

Our critique distinguishes Rytina's findings from his conclusions. In our opinion, the findings are correct and the conclusions are all wrong. Rytina obtains large intergenerational correlations in the 1972–86 GSS because the canonical scores depend on a very large number of parameters. Analytically, this alone accounts for almost all of the difference between Rytina's intergenerational correlations and those based on the SEI. We also used Rytina's scores to cross-validate Rytina's empirical findings in

one or more of three bodies of data, the 1987–90 GSS (Davis and Smith 1990), the 1962 Occupational Changes in a Generation Survey (OCG), and the 1973 OCG (Featherman and Hauser 1983). The cross-validations do not support Rytina's findings: (1) intergenerational correlations of occupational standing are about the same or smaller when Rytina's scales are used as when the Duncan SEI is used; (2) the direct intergenerational effect of father's occupational standing—the effect that is not mediated by educational attainment—is smaller than the share estimated by Rytina in the 1972–86 GSS; (3) within levels of educational attainment, intergenerational correlations of occupational standing are about the same when the SSIC is used as when the Duncan SEI is used, that is, they are negligible. Since we cannot confirm Rytina's findings in other samples, we cannot agree with his conclusion that the empirical evidence supports a preference for the SSIC over the Duncan SEI.

There are obvious practical limits to the value of the SSIC in research. One of the advantages of other widely used scales is that the relevance and validity of the criteria on which they are based, for example, popular ratings of social standing or occupational education and income, are widely acknowledged. We doubt that there will be similar demand for an occupational scale that maximizes the intergenerational correlation between occupations. This is important because there is scientific value in using the same occupational scale in different fields and disciplines, for example, in studies of political behavior, educational performance, and mortality, as well as in studies of social mobility.

Also, despite Rytina's efforts to enlarge the number of occupation lines in his analysis, the SSIC ignores about one-quarter of the census occupation titles and about one-tenth of the work force. We do not believe that it would be acceptable to the research community to censor a substantial number of occupations or workers in this way. For example, Duncan's (1961, pp. 110–13) SEI was developed and accepted in large part because, for the first time, it provided an index of the status of *all* U.S. occupations; it took another decade to develop a prestige scale for all occupations (Siegel 1971). It would be possible to cover all occupations with an index constructed in the same fashion as the SSIC by adopting an intermediate-level aggregation of occupation lines rather than using all detailed lines. However, this could only be carried out with a corresponding loss of occupational detail.

We have also assessed the conceptual basis for Rytina's claim that the SSIC (or a similar scale) provides a superior framework for the measurement of intergenerational occupational persistence, and we find none. While we think there is a place in mobility research for canonical correlation, association models, and related criterion scoring methods, we do not believe there is any reason to replace scales based on measured occu-

pational characteristics with the SSIC (or a similar scale).[3] Even if the empirical evidence showed that correlations based on the SSIC (or a similar scale) were consistently larger than those based on conventional scales, we would not accept the SSIC as a new basis for mobility studies. On the contrary, we think that there should be more basic research on measurable dimensions of occupations that persist across generations.

CORRECTING THE CANONICAL CORRELATIONS

Rytina acknowledges that his analysis is based on a single sample and that replication in other samples is desirable. He also acknowledges the possibility that the scale values he obtained may be affected by overfitting chance fluctuations in the sample data. Unfortunately, rather than carrying out a true cross-validation, he relies upon various tests internal to the 1972–86 GSS sample, for example, a comparison of the intergenerational SSIC correlation between populous and sparse occupational lines.

The key problem is that Rytina ignores the sampling variability of his intergenerational correlations.[4] When the SSIC is estimated, the intergenerational correlations increase (relative to conventional scales or indexes) for two reasons: first, the SSIC may reflect dimensions of occupations that persist, but are not correlated with the conventional measures, and, second, the SSIC capitalizes on chance associations between occupations in the original sample.[5] In effect, when he estimates the canonical correlation between father's occupation and offspring's occupation, Rytina is computing a multiple correlation based on 614 independent variables. When he constrains the scales to be the same for offspring as for fathers, he is still estimating an equation with 307 independent variables. Even in a sample of nearly 8,000 cases, the effect of this profligate use of

[3] However, association models, which have been extended to multidimensional classifications, provide a superior methodological basis for empirically based rankings of occupational categories, comparisons of those rankings across populations, and comparisons of the openness of mobility regimes. Goodman (1984) provides the key exposition of these models and methods, including a comparative analysis of association models and canonical correlation models.

[4] In the one place (1992, table 6) where Rytina reports standard errors of correlations based on the SSIC, they are apparently based on the naive and erroneous assumption that the SSIC scores were not estimated in the same data used to estimate the correlations.

[5] There is, or was, a similar problem in the initial construction of the Duncan SEI, which is a weighted average of aggregate measures of the educational attainment and income of men in the 1950 census This gave rise to criticisms that the correlations between the SEI and education or income were built in, which is true, and that they were artifacts, which is not true (Blau and Duncan 1967, pp. 117–28). In any event, all of the analyses reported herein are far removed from the data that were used to construct the SEI.

degrees of freedom cannot be ignored; the canonical correlation is much larger in the sample than we would expect it to be in the population.

Rytina estimates a canonical correlation of 0.489 between father's occupation and offspring's occupation in 7,965 cases from the 1972–86 GSS. Using one of Lawley's (1959) formulas, we estimate a corrected correlation of 0.411 when the sample canonical correlation is 0.489 in a simple random sample of 7,965. But the GSS sample is actually much less efficient than a simple random sample; it is approximately two-thirds as efficient (Davis and Smith 1991). Thus, the problem of statistical unreliability is greater than would be suggested by the size of the GSS sample.[6] If we apply Lawley's formula again, assuming a simple random sample two-thirds the size of the GSS, we estimate the corrected correlation, 0.337. This is virtually the same as the observed intergenerational SEI correlation, 0.338.[7]

Thus, the "dramatic" difference between intergenerational SEI correlations and intergenerational canonical correlations is just about what we should expect if the true intergenerational correlation were that observed for the SEI, given only the size and design of the GSS sample and the number of parameters to be estimated. Rytina's findings of greater ascription are essentially an artifact of his failure to take account of the sample-population distinction.

INTERGENERATIONAL CORRELATIONS OF OCCUPATIONAL STANDING

In our cross-validations, we compare the behavior from sample to sample of intergenerational correlations based on Rytina's canonical scales with that of correlations based on the NORC prestige scale and the Duncan SEI. The first column of table 1 reproduces the correlations from which Rytina concludes that the NORC prestige scale and the SEI are less highly correlated across generations than the rough and ready (hereafter referred to as R&R), the SSIC, or the canonical scales in which father's

---

[6] This problem also may have a significant effect on Rytina's identification of nonlinearities in correlations based on the SEI. He identifies 47 occupation lines for which nonlinearities are nominally significant with $P < 05$, but if the GSS is two-thirds as efficient as a simple random sample of the same size, the true probability level exceeds .10. We think this is a loose criterion for the identification of outliers

[7] We are grateful to Steve Rytina for correcting an error in our application of a related formula in Lawley (1959). Incidentally, it is unnecessary to correct SEI correlations for loss of degrees of freedom, as we have corrected Rytina's canonical correlations, because the SEI uses only one degree of freedom.

TABLE 1

INTERGENERATIONAL CORRELATIONS OF OCCUPATIONAL STANDING

	GENERAL SOCIAL SURVEY			20 Outliers Deleted		OCCUPATIONAL CHANGES IN A GENERATION, 1973		20 Outliers Deleted	
	308 Occ, 1972–86	306 Occ, 1972–86	306 Occ, 1987–90	1972–86	1987–90	306 Occ, Current	306 Occ, First	Current	First
Prestige	.256	.256	261	217	.227	.263	.294	218	266
SEI	.338	348	.301	.319	.274	.375	.423	.343	.365
R&R	427	.430	267	.390	.249	.	.	. .	.
SSIC	.450	453	.267	.412	.261	.376	.474	.349	.369
Canonical	.489	.493	220	.454	.208	.320	386	.294	.298
Ratios of correlations:									
R&R/SEI	1 26	1 24	89	1 22	.91	. .	1 12	. .	.
SSIC/SEI	1.33	1 30	.89	1.29	.95	1.00	.91	1.02	1 01
Canonical/SEI	1.45	1.42	.73	1.42	.76	.85	.91	86	.82
N	7,965	7,777	2,602	7,355	2,455	23,016	21,202	20,559	17,694

NOTE —All findings except those in the first column or those including NORC prestige scores exclude military occupations, in the 1972–86 GSS, there are 7,777 cases after fathers or offspring with military occupations have been excluded Occ = occupations.

and offspring's occupations are permitted to have distinct scores.[8] In the second column of table 1, we show similar correlations, based upon a slightly different use of the 1972–86 GSS sample. First, we have deleted the two military lines. There is no reason to believe that the SEI scores for those lines accurately represent current members of the armed forces; the data used by Duncan (1961, pp. 263–75) pertained to former members of the armed forces who were unemployed at the time of the 1950 census. Also, there are no directly measured NORC prestige scores for members of the armed forces, and it is conventional to limit analyses of social mobility to the civilian noninstitutional population. Second, we have corrected the NORC prestige score for carpenter's helpers (750) from "0," as listed incorrectly in the source used by Rytina, to the correct value of "23." These changes barely affect the correlations.

In the last three rows of table 1, we report the ratios of correlations based on Rytina's three scales to the correlations based on the SEI. We believe these ratios provide an appropriate basis for comparing the performance of the four scales. We see no reason to dwell on the lower intergenerational correlations of prestige scores (Jencks 1990). The comparison with correlations based on the SEI takes account of the variation from sample to sample in the intergenerational correlation of occupational socioeconomic status. Both in the first and second columns of table 1, each of Rytina's scales has a substantially higher intergenerational correlation than the SEI, and the differential increases from the scale that is partly dependent on the SEI—the R&R score—to the unconstrained canonical scoring. Correlations based on Rytina's scales are 25%–45% larger than those based on the SEI.

In the third column of table 1, we report intergenerational correlations in an independent sample, the 1987–90 GSS. The population definition and codings of all variables are identical to those in the 1972–86 GSS; only the years and the random sampling outcomes differ. The findings are very different in this sample. While the prestige correlations are virtually identical, the other three correlations are smaller. The SEI correlation declines by 13.5%, from .348 to .301, suggesting that there has been a decline in occupational status persistence in the United States. However, the R&R correlation declines by 37.9%, from .430 to .267, and the SSIC correlation declines by 41.1%, from .453 to .267, so it is smaller than the SEI correlation and about as large as the NORC prestige correlation. The correlation of canonical scores declines even more, by 55.4%, from .493 to .220. Rytina's three scales yield correlations that are, respectively,

[8] Rytina did not give us his R&R scale values, and we reconstructed them from the description in his text. We have reproduced the findings in his table 1 using our version of the R&R scale, and we are satisfied that it is the same scale he constructed.

89%, 89%, and 73% as large as those based on the SEI in the 1987–90 GSS.

Notably, the correlation based on the R&R scale shrinks almost as much as that based on the SSIC. This is significant because the R&R scale is the result of a single iteration of one method for estimating a symmetrically scaled canonical correlation.[9] It is not overfitting in the later stages of estimation that accounts for sample-to-sample shrinkage of the intergenerational correlations, but a failure to replicate the most obvious nonlinearities in the intergenerational relationship of SEI scores. That is, Rytina claims "superior validity" for the SSIC because intergenerational correlations increase when he smooths out nonlinearities in the relationship between father's and offspring's SEI; the claim fails because the nonlinearities are not replicable. When we use Rytina's scales in the 1987–90 GSS, there is no longer any indication that they yield larger intergenerational correlations than the SEI.[10]

To be sure, one could obtain higher intergenerational correlations in the 1987–90 GSS sample by estimating another set of constrained and unconstrained canonical scores, but these correlations would be subject to exactly the same problem of overfitting as those in the 1972–86 sample. Furthermore, just as the SSIC correlations shrink in the 1987–90 GSS sample, we expect that correlations based on canonical scores in the 1987–90 GSS will shrink in other samples. The problem is that we want a measure of occupational standing that can be used to assess differences in intergenerational correlations from one population to another, but the SSIC cannot help us to do this as long as one of the samples of interest is the one from which it was estimated. And outside that sample, the SSIC loses its seeming advantage.[11]

We agree with Rytina that statistical outliers—extreme values in the distribution of the SSIC—may be responsible for the unreliability of correlations based upon it. For this reason, we repeated our analysis of the 1972–86 and the 1987–90 GSS samples, after we eliminated 20 occupation lines falling more than 2.5 standard deviations (SDs) from the mean of the SSIC. As shown in column 4 of table 1, these exclusions have little effect on the findings in the 1972–86 data. However, the 1987–90 GSS data again show no change in the NORC prestige correla-

[9] See Rytina (in this vol.), n. 14.

[10] We have also cross-validated Rytina's scales separately for men and women in the 1972–86 and in the 1987–90 GSS, and we have looked at the behavior of Rytina's scales in relation to that of more recent versions of the Duncan SEI (Stevens and Featherman 1981). All of these analyses yield essentially the same findings as those reported herein.

[11] The proper way to deal with this problem would be to test for changes between populations in canonical scores and in correlations based upon them.

tion, a small decline in the SEI correlation, and very large declines in the correlations based on the SSIC, R&R, and canonical scores. Evidently, the problem of sample-to-sample shrinkage in correlations based on Rytina's scales cannot be attributed to this small set of extreme values on the SSIC.

In the last four columns of table 1, we report cross-validations of intergenerational correlations of Rytina's SSIC and canonical scales in the 1973 Occupational Changes in a Generation Survey (Featherman and Hauser 1978). The data pertain to intergenerational correlations of father's occupation with son's current occupation and with son's first occupation among men 25–64 years old who were in the experienced civilian labor force in March 1973. Again, the analysis is restricted to the 306 civilian occupation lines that were included in Rytina's analysis. Here, the findings are more favorable to the SSIC than in the 1987–90 GSS sample. The intergenerational correlations of SEI and SSIC scores are almost the same, except the SSIC correlation is slightly larger (0.474 vs. 0.423) in the case of father's occupation and son's first occupation; this differential disappears when the 20 outliers are removed. The intergenerational correlations of the canonical scales are consistently lower than those of either the SSIC or the SEI; we think this is clear evidence of shrinkage due to overfitting in the 1972–86 GSS data. As is well known, the intergenerational correlations of the NORC prestige scores are lower than those of the SEI and, consequently, also lower than those of Rytina's scales. Finally, deletion of the 20 outliers lowers all of the correlations slightly, but—excepting the negligible difference between SEI and SSIC correlations—the general pattern of findings remains the same. From this second cross-validation, we again find no reason to conclude that correlations based on the SSIC will be larger than those based on the SEI, outside of the sample in which the SSIC was estimated.

EDUCATIONAL ATTAINMENT AS AN INTERVENING VARIABLE

Rytina argues that his scales yield larger estimates of the direct effect of father's occupational standing on offspring's occupational standing using a simple three-variable causal model: father's occupational standing affects educational attainment, and both of those variables affect offspring's occupational standing. As shown in the first two columns of table 2, Rytina's scales lead to estimates of the direct intergenerational effect that are 1.9–2.7 times larger than estimates based on the SEI. Table 3 contains evidence that his scales also lead to much larger estimates of the share of the intergenerational effect that is direct: 53.4% for R&R, 58.8% for SSIC, and 67.3% for the canonical scores, compared to only 36.3% for the SEI.

TABLE 2

DIRECT EFFECTS OF FATHER'S OCCUPATIONAL STANDING ON OFFSPRING'S OCCUPATIONAL STANDING

| | GENERAL SOCIAL SURVEY | | | | | OCCUPATIONAL CHANGES IN A GENERATION, 1973 | | | |
| | | | | 20 Outliers Deleted | | | | 20 Outliers Deleted | |
	308 Occ, 1972–86	306 Occ, 1972–86	306 Occ, 1987–90	1972–86	1987–90	306 Occ, Current	306 Occ, First	Current	First
Prestige	.080	.080	.091	.068	.084	.097	.118	.094	.121
SEI	123	124	.110	113	104	149	.181	.140	.149
R&R	.228	.231	092	207	.101	...	...	...	...
SSIC	.265	.269	097	.236	115	.174	.265	.169	.187
Canonical	.329	.333	080	.304	093	.146	.195	.144	148
Ratios of effects:									
R&R/SEI	1.86	1.87	84	1.83	.97	...	...	...	...
SSIC/SEI	2 16	2 17	.88	2 09	1 10	1.17	1 46	1 21	1.25
Canonical/SEI	2 68	2.69	72	2.69	.89	98	1.07	1.03	1.00
N	7,965	7,777	2,602	7,355	2,455	23,016	21,202	20,559	17,694

NOTE.—All findings except those in the first column or those including NORC prestige scores exclude military occupations, in the 1972–86 GSS, there are 7,777 cases after fathers or offspring with military occupations have been excluded

TABLE 3

PERCENTAGE OF THE CORRELATION BETWEEN FATHER'S AND OFFSPRING'S OCCUPATIONAL STANDING THAT IS NOT MEDIATED BY EDUCATIONAL ATTAINMENT

	GENERAL SOCIAL SURVEY			20 Outliers Deleted		OCCUPATIONAL CHANGES IN A GENERATION, 1973		20 Outliers Deleted	
	308 Occ, 1972–86	306 Occ, 1972–86	306 Occ, 1987–90	1972–86	1987–90	306 Occ, Current	306 Occ, First	Current	First
Prestige	31.4	31 3	34.8	31 2	36 9	37.1	40.1	42.9	45 4
SEI	36.3	35.6	36.5	35 4	37.9	39.7	42 8	40.7	40.8
R&R	53 4	53.7	34 6	53 1	40 7	.	.	.	.
SSIC ..	58.8	59.3	36 4	57.4	44.0	46.2	55.9	48.3	50.7
Canonical	67.3	67 7	36.3	67.0	44 6	45 7	50 4	48.9	49.8
Ratios of percentages:									
R&R/SEI	1 47	1.51	.95	1 50	1 07	.	.	.	.
SSIC/SEI	1.62	1 67	1.00	1.62	1.16	1.16	1.31	1 19	1.24
Canonical/SEI	1.85	1.90	1.00	1 89	1.18	1 15	1 18	1 20	1 22
N	7,965	7,777	2,602	7,355	2,455	23,016	21,202	20,559	17,694

NOTE.—All findings except those in the first column or those including NORC prestige scores exclude military occupations, in the 1972–86 GSS there are 7,777 cases after fathers or offspring with military occupations have been excluded

   In our opinion, Rytina invites the reader to assess the role of education as an intervening variable in the stratification process by reference *only* to the effects of father's occupational status in the 1972–86 GSS, and he emphasizes the revisionist implications of his findings in those data. But his findings are not a great deal different from those of Blau and Duncan (1967, pp. 169–70) that are based on the Duncan SEI. They estimated a correlation of .405 between father's SEI and son's current SEI and a correlation of .417 between father's SEI and the SEI of the son's first job. These two estimates—in an earlier sample, to be sure—are not much smaller than those estimated by Rytina using the R&R and SSIC in the 1972–86 GSS. Using the original Blau-Duncan data, and applying Rytina's three-variable model, 44.0% of the intergenerational effect on son's current SEI is direct and 53.8% of the intergenerational effect on the SEI of son's first occupation is direct. Both of these are substantially larger than Rytina's estimate of 36.3% for SEI of current occupation in the 1972–86 GSS data, and the latter, at least, is comparable to Rytina's estimates based on the R&R and SSIC.

   Where, then, is the disagreement between Rytina's finding and the conventional view? The problem is that Rytina looks only at the effect of father's occupational standing on offspring's occupational standing and not, more broadly, at the effect of social background on occupational standing. Even by the standards of a generation ago, Rytina's three-variable model cannot be regarded as a plausible or complete representation of the stratification process. For example, the original Blau-Duncan model included father's education as well as father's occupational standing among the exogenous variables. In the original Blau-Duncan model, 55.2% of the effect of father's occupational SEI on son's current SEI is direct, and 64.6% of the effect of father's occupational SEI on the SEI of son's first occupation is direct.[12] However, in that same model, *none* of the effect of father's education on the SEI of current or of first occupation is direct; all of the effect of paternal education on son's occupational attainment is mediated by schooling. Subsequent models of status attainment typically include other social background variables, like number of siblings, farm origin, southern birth, race, and intact family. The general finding that educational attainment is a key intervening variable in the stratification process is not based merely upon relations between the occupational ranks of parents and offspring, but upon a more extensive set of findings about the effects of several social background characteristics. For example, in Featherman and Hauser's (1978, p. 274) regressions of the SEI of first job on social and economic background characteristics among birth cohorts reaching age 16 between 1925 and 1965, the direct

---

[12] These are percentages of reduced-form coefficients, not of correlations.

effect of father's occupational SEI averages 42% of its total effect, while the direct effect of number of siblings averages 34.5% of its total effect; the direct effect of father's education averages 3.5% of its total effect; and the direct effect of broken family averages 5.6% of its total effect.[13] That is, there is no substantial basis for the claim that education mediates more than about half of the effect of father's occupational standing on offspring's occupational standing, but there is a good deal of evidence of a stronger mediating role in the case of other background variables.

Despite the weak rationale for Rytina's analysis of direct and indirect effects, we have cross-validated it in the 1987–90 GSS and in the 1973 OCG surveys. The third column of table 2 gives our estimates of the direct intergenerational effects of father's occupational standing that are obtained when Rytina's three-variable model is estimated in the 1987–90 GSS. Again, Rytina's major finding disappears in the 1987–90 data. The direct effects of the R&R, SSIC, and canonical scales are somewhat smaller than those of the SEI and similar to that of the prestige scale. The direct effect of the SEI is almost 90% as large in 1987–90 as in 1972–86, but the direct effect of the SSIC is only 36% as large in 1987–90 as in 1972–86. As shown in table 3, in the 1987–90 GSS, the direct effects of father's occupational standing account for about one-third of its correlation with offspring's standing, regardless of the scale that is used. The fourth and fifth columns of tables 2 and 3 repeat the comparison between the 1972–86 and 1987–90 GSS surveys, but with 20 outlying occupations removed; again, Rytina's findings almost disappear in the 1987–90 GSS. In the last four columns of tables 2 and 3, we report cross-validations in analyses of mobility to current occupations and to first occupations in the 1973 OCG survey. Here, the SSIC, but not the canonical scale, yields larger direct effects than the SEI. We do not know why the SSIC has larger direct effects in the OCG data; it may be the closer temporal proximity of the 1973 OCG to the period covered by the 1972–86 GSS.[14] Still, taking into account the findings in both the GSS and the 1973 OCG, we are not convinced that analyses based on the SSIC will yield evidence of greater direct ascription than those based on the SEI.

ASCRIPTION WITHIN LEVELS OF SCHOOLING

In addition to estimating the direct effect of father's occupational standing in the general population, Rytina looks at the effect within six specific levels of completed schooling: 0–7 years, 8 years, 9–11 years, 12 years,

[13] See n. 11 above.
[14] See the discussion of table 5 below.

13–15 years, and 16 or more years. His analysis addresses and claims to disconfirm Hout's (1988) finding that the effect of father's occupational status on offspring's status is negligible among college graduates.[15] Our slightly modified version of Rytina's analysis, shown in the first of each pair of columns in table 4, reproduces and extends his findings. In the 1972–86 GSS, within each level of completed schooling, intergenerational correlations of occupational standing based on the SSIC and canonical scales greatly exceed those based on the SEI or prestige scales. The disparity in findings is indeed "dramatic" among college graduates, the locus of Hout's analysis. The intergenerational correlation of the SEI is only .072, while those of the SSIC and canonical scales are .288 and .359.

The second of each pair of columns in table 4 reports corresponding correlations estimated in the 1987—90 GSS. In every case the decline in correlations based on Rytina's scales is as dramatic as the previously noted disparity in findings within the 1972–86 GSS. For example, among college graduates, the intergenerational correlation of the SSIC falls from .288 to .097, while that of the SEI falls from .072 to .050. We think that both correlations in the 1987–90 data are negligible by any reasonable standard. Overall, while there is variation in the correlations from one level of schooling to the next, we see no evidence that any of those correlations are more than modest in size or that those based on Rytina's canonical scales are larger than those based on the SEI or NORC prestige scales.[16] As in the case of his first two findings, Rytina's third major empirical finding fails the test of cross-validation. There is no reason to conclude from Rytina's analysis that SEI correlations underestimate intergenerational status ascription.

THE USE AND ABUSE OF CANONICAL CORRELATIONS

Apart from the empirical issue of cross-validation, there is an important question about the meaning of the SSIC and the intention of Rytina's proposal. Would use of the SSIC be justified if there were no evident

---

[15] The SEs that Rytina reports in his table 6 for the SSIC are incorrect. He apparently has tried to use a standard formula for the standard error of a zero-order correlation, but this standard formula does not take into account the degrees of freedom lost in estimation of the SSIC. That is, Rytina underestimates the sampling variability of correlations based on the SSIC in the 1972–86 GSS data. Thus, Rytina's reports of t-statistics and probability values associated with the SSIC are also incorrect. Also, Rytina reports 1,847 as the sample count in two education subgroups, leading to an inconsistency between the counts in the table and the total number of observations in the sample.

[16] We have repeated these analyses after removing 20 extreme observations on the SSIC, and the findings are the same.

TABLE 4

INTERGENERATIONAL CORRELATIONS OF OCCUPATIONAL STANDING BY LEVEL OF COMPLETED SCHOOLING

	0–7 Years		8 Years		9–11 Years		12 Years		13–15 Years		16 or More Years	
	1972–86	1987–90	1972–86	1987–90	1972–86	1987–90	1972–86	1987–90	1972–86	1987–90	1972–86	1987–90
Prestige	.111	−.006	.186	007	.119	.160	.113	.123	.062	.091	.042	.071
SEI	137	.272	.082	−.031	135	157	.170	159	.145	116	.072	.050
SSIC	331	114	279	058	276	134	.299	138	.310	076	.288	.097
Canonical	.402	−.017	.343	.036	.302	123	377	119	356	065	.359	081
N .... :	341	61	331	46	1,059	251	2,696	826	1,567	688	1,783	730

problems of cross-validation? Answering this question requires an understanding of what the SSIC is supposed to measure and whether this is in fact the same thing that is supposed to be measured by the occupational scales used in models of status attainment.

Rytina says he intends to improve the measurement of the "intergenerational continuity of occupation." What is "occupation" (in the singular) and in what sense is it continuous across generations? An occupation is a socially recognized collection of similar jobs, where similarity is usually defined with respect to activities and duties performed, goods or services produced, forms of employment tenure, or—in broader classifications—educational requirements and economic rewards. As such, occupation is not a quantitative variable, and we can only assign numeric values to occupations in relation to one or more occupational characteristics.

In previous mobility studies continuity has been studied in two different respects. First, whether the child follows the same occupation as the father is treated as a question of continuity of occupations (plural). This is an essential definition of continuity in discrete models and corresponds to a concern with the proportion of cases that fall on the main diagonal of a mobility table. Unfortunately, the proportion of cases that is continuous in this sense depends on the level of detail of the occupational categorization and on the distribution of persons across occupations in each generation, as well as on specific mechanisms of occupational persistence or mobility.

A second mode of studying continuity has been achieved by refining the question and concentrating on only one or a few measurable properties of all occupations. With measures of such properties as prestige, education, or income it is possible to ask how much the offspring's occupation is like or unlike the parent's, even when the two are not in the same category. The degree of likeness or difference between the generations, in such an approach, has a specific metric corresponding to the units of the measured property. This refinement of the problem of continuity of occupations to become the problem of the continuity of properties associated with occupations is the essential shift in perspective in linear status attainment models, as compared with traditional models for mobility tables.[17]

Of course, the earlier focus on the continuity between generations in occupational categories as such did not imply a disregard for the properties of the occupations in the categories. Instead, it was implicit that

---

[17] However, recent developments of log-linear and log-multiplicative models have made it possible to combine the two perspectives (Goodman 1979, Hout 1984a; Hout 1984b; Hout 1988).

occupational categories had stable and relatively homogeneous proper-
ties, and this stability and homogeneity made it worth inquiring into
continuity across generations.

Rytina's SSIC, however, deliberately disregards the existential, mea-
surable aspects of life experienced in occupations. The SSIC places occu-
pations close to one another purely on the basis of relative rates of inter-
generational movement. To Rytina, an occupation that characteristically
has better prospects for the next generation, say, in terms of SEI, should
be ranked higher than one that has the same SEI but poorer next-
generation prospects. Though SSIC manifestly does not take the status
of occupations into consideration at all, Rytina justifies its use in part on
the grounds that it corrects the SEI by repositioning occupations in closer
correspondence with "the details of ascent and descent."

The idea that SEI values of occupations should be corrected by ad-
justing them on the basis of average intergenerational moves from and
to higher and lower occupations is misguided. Though the precise mean-
ing of SEI may be unclear, an undoubted advantage of the measure,
which it shares with prestige, is that it is determined independently of
the intergenerational movements that it is intended to gauge and predict.
Socioeconomic status and prestige are properties of occupations, not
properties of the relative standings between occupations of one generation
and those of its offspring or parents. If two occupations have the same
SEI but the first shows higher average SEI for its offspring, it may seem
reasonable to raise its ranking to reflect the good fortune of its children.
But this good fortune is not enjoyed by the parents and should not be
imputed to them. On the other hand, if two occupations share the same
SEI but the first has higher average levels of SEI among its parents, it
does not seem reasonable to raise the first occupation's ranking, since
there is no obvious reason why downwardly mobile offspring should be
judged better off than their stable or upwardly mobile peers. Yet Rytina's
procedure does give downwardly mobile occupations a higher ranking,
and, furthermore, it adjusts an occupation's ranking by the same amount
whether it is downward mobility from the preceding generation or up-
ward mobility to the next that is involved.

We believe, in contrast, that an average decline or rise in SEI between
generations is a social fact, which, if true, should be revealed by the
measures used in mobility studies, rather than intentionally obscured by
them. The study of social mobility should, to the greatest degree possible,
be a study of the measurable properties of individuals, including the
properties of their occupations, and the means by which individuals come
to have those properties. We should study status, prestige, income,
wealth, health, and well-being, but not a measure that places occupations

close to one another solely by relative rates of interchange. Our empirical and theoretical interest lies in the degree to which relative rates of interchange do or do not correspond to stability in measurable properties of occupational roles.

This said, it is still true that canonical correlation and similar procedures have their place. The canonical correlation provides a useful reference point for judging the degree to which a single scale or set of scales based on measured properties of occupations can account for the observed pattern of intergenerational movement. This is the use that Klatzky and Hodge (1971) made of the method in their exemplary analysis of Blau and Duncan's (1967) 17-category mobility table. Similarly, Hout's (1984a) status, autonomy, and training (SAT) model of the mobility table might be compared with log-linear or log-multiplicative models that generate intrinsic scalings of occupational categories. It is quite a long step from such uses to substituting canonical scores for measured properties of occupations.

A second, proper use of canonical scaling and related procedures is to determine the number of linear dimensions required to account for the association in a mobility table. For example, Klatzky and Hodge (1971) identified two distinct dimensions of occupational mobility in Blau and Duncan's (1967) 17-category mobility table, and similar tests for dimensionality can be constructed within the framework of log-linear and log-multiplicative models.

One seeming advantage of scales based on log-linear and log-multiplicative models is that they are marginally invariant, while canonical scales and correlations are not. However, it is possible to sidestep this problem. Hauser and Featherman (1977) and Featherman and Hauser (1978) estimated canonical correlations in a series of mobility tables that were standardized to fit a common set of relative mobility chances and thus demonstrated the effects of trends in occupational distributions on intergenerational occupational correlations. These trends (toward lower intergenerational correlations) closely followed the trends in correlations in the observed tables.

However, because of the problem of shrinkage, great caution should be exercised in any canonical analysis of mobility classifications, especially if the samples are small and the number of occupations is large. Even at the high levels of aggregation common in most studies of mobility tables, canonical scores exhibit shrinkage when they are applied to independent samples. Table 5 shows this effect in the two GSS samples and in Blau and Duncan's 1962 OCG data (1967, p. 496, table J2.1), using the standard 17-category classification. The rows of table 5 identify scales for the categories according to their sources, while the columns identify the

TABLE 5

FATHER-OFFSPRING CORRELATIONS OF 17-CATEGORY SEI, CANONICAL SCORES, AND
AVERAGED CANONICAL SCORES IN THREE U.S. SAMPLES

	SAMPLE, CORRELATION, AND SHRINKAGE		
	OCG 1962	GSS 1972–86	GSS 1987–90
SEI	403	338	305
Scores based on 1962 OCG			
EIG1	.449	364	.301
Shrinkage	1.000	967	886
Average EIG1	.445	.368	297
Shrinkage	1.000	.986	.882
Scores based on 1972 to 1986 GSS:			
EIG1	.427	.383	296
Shrinkage	.935	1.000	856
Average EIG1	437	374	296
Shrinkage	.980	1 000	.877
Scores based on 1987 to 1990 GSS:			
EIG1	.371	324	.337
Shrinkage	.833	868	1 000
Average EIG1	.368	330	.326
Shrinkage	.854	.913	1 000

NOTE —Rows labeled EIG1 give the father-offspring sample correlations of EIGf and EIGo from the specified sources. Rows labeled "Average EIG1" replace EIGf and EIGo with their arithmetic averages SEI is the average Duncan score within each of the 17 categories Shrinkage is the ratio of the correlation of a canonical score to the correlation of the SEI in a given sample and year relative to the corresponding ratio in the sample and year in which the canonical scores were estimated

samples to which the scales are separately applied. The entries in the table are the correlations between fathers' and offsprings' occupations obtained with each combination of a scale and sample.

Row 1 of table 5 gives the intergenerational correlation of SEI at the 17-category level of aggregation. The temporal decline in this correlation again suggests a decrease in the intergenerational association of occupational standing over time. Row 2, EIG1, shows the correlations observed in the samples using canonical scores derived from the 1962 OCG data. The value .449 in column 1 of row 2 is the canonical correlation in the 1962 data, while the values .364 and .301 in columns 2 and 3 are intergenerational correlations in the GSS samples and are based on the canonical scores estimated in the 1962 data. Row 4, average EIG1, gives the same information using a simple arithmetic average of the fathers' and sons' canonical scores in the 1962 data. Because the SEI correlation declines over time, we have normed our comparisons of canonical correlations with SEI correlations on the ratios of canonical correlations to those of the SEI. Thus, in rows 3 and 5, both labeled "shrinkage," we

show the ratio of the canonical correlation to the SEI correlation in each sample to the corresponding ratio in the (1962 OCG) sample in which the canonical scores were estimated. In each of the GSS samples, the intergenerational correlation of canonical scores shrinks relative to the correlation based on the SEI. Table 5 also reports similar analyses of canonical scores estimated in the 1972–86 GSS and in the 1987–90 GSS. In each case, there is shrinkage in the correlation of canonical scores estimated in another sample, and the shrinkage is larger than we would expect from sampling variability alone.[18] Even at the high level of aggregation employed by Klatzky and Hodge (1971), canonical weighting cannot be relied upon to yield scores permitting valid, direct comparisons of intergenerational occupational persistence across populations. Even a modest expansion of occupational detail will yield substantial differences between sample values of the canonical correlation and estimates adjusted for degrees of freedom. For example, if the true correlation lies between 0.4 and 0.5, and we use a 100-category occupational classification, Lawley's correction leads us to expect sample estimates from 6% to 10% higher than the true intergenerational correlation in a sample like the 1972–86 GSS.

If canonical scores or similar scores are used as baseline associations in tabular models seeking to describe nonlinear patterns of mobility, the shrinkage of association when using a canonical score in a new sample may be inconsequential, since new scores will be extracted in each sample, and exact score values in any single sample will be of secondary interest. However, Rytina's proposal that constrained canonical scores be taken as measures of social standing strongly suggests that they should be used in more than one sample. Only by using the scores in more than a single sample can questions of cross-sample or cross-time variability in the intergenerational returns to social standing be addressed. Reestimating an SSIC in every new sample would confound changes in canonical scores, changes in occupational distributions, and sampling variability.[19]

However, in the present instance, no matter how the data are scaled, there does appear to be a decline in the intergenerational correlation across the three surveys. The decline appears in the SEI correlations, in the unconstrained canonical correlations (.449, .383, and .337), and in the correlations of average canonical scores (.445, .374, .326). Based upon the work of Featherman and Hauser (1978), we suspect that changing

[18] Lawley's (1959) formula implies that the sample canonical correlation should be about 98.4% of the sample correlation in the 1972–86 GSS when the analysis is based on 17 occupational categories.

[19] Again, there are appropriate methods to address these problems in the case of log-linear and log-multiplicative models for ordered categories (Goodman 1984).

occupational distributions, as well as changes in the association between father's and offspring's occupation (Hout 1988) are sources of these trends.

CONCLUSION

In the context of linear models of the stratification process, Rytina's proposed use of canonical scores must be judged wanting, in comparison with the well-established use of measured properties of occupations, such as NORC prestige scores or the Duncan SEI. Although we have used the SEI as a standard in our evaluation of Rytina's proposals, we do not want to advocate its use too strongly.[20] The most that can be said for it is that the SEI does appear to represent several of the sources of intergenerational occupational persistence and that future uses of it will be comparable to past uses. At the same time, we believe that the study of social stratification would be well served by the development of models that disaggregate occupational education from occupational income and that add other measurable occupational characteristics to models of intergenerational occupational stratification.

In models of mobility tables, canonical and related scaling models can be justified as an economical means of extracting linear associations from the table to evaluate nonlinear, categorical effects, to assess the dimensionality of association, and—with better statistical methods than Rytina has used—to provide a baseline against which substantive models of occupational mobility may be assessed. The ultimate goal in the analysis of mobility tables should also be to model relationships in measurable properties of occupations and persons. Only such models can give substance to our measurements of the intergenerational persistence of occupational roles.

REFERENCES

Blau, Peter M , and Otis Dudley Duncan. 1967. *The American Occupational Structure*. New York. Wiley.
Davis, James A., and Tom W. Smith. 1990 *General Social Surveys, 1972 to 1990* (MRDF). NORC ed Chicago. NORC Distributed by the Roper Center for Public Opinion Research, Storrs, Conn.
———. 1991. *The NORC General Social Survey: A User's Guide*. Chicago: NORC.
Duncan, Otis Dudley. 1961. "A Socioeconomic Index for All Occupations." Pp 109–38 in *Occupations and Social Status*, edited by Albert J. Reiss, Jr. New York: Free Press.
Featherman, David L., and Robert M. Hauser. 1978. *Opportunity and Change* New York: Academic.

[20] Hodge (1981) offers an insightful critique of the Duncan SEI.

————. 1983. *Occupational Changes in a Generation, 1973* (MRDF). DPLS ed. Washington, D.C.: U.S. Bureau of the Census. Distributed by Data and Program Library Service, University of Wisconsin, Madison.

Goodman, Leo A. 1979. "Simple Models for the Analysis of Association in Cross-Classifications Having Ordered Categories." *Journal of the American Statistical Association* 74 (September)· 537–52

————. 1984. *The Analysis of Cross-Classified Data Having Ordered Categories*. Cambridge, Mass.: Harvard University Press.

Hauser, Robert M , and David L. Featherman. 1977. *The Process of Stratification: Trends and Analyses*. New York: Academic.

Hodge, Robert W. 1981. "The Measurement of Occupational Status " *Social Science Research* 10 (December): 396–415.

Hout, Michael. 1984*a*. "Status, Autonomy, and Training in Occupational Mobility." *American Journal of Sociology* 89 (May). 1379–1409

————. 1984*b*. "Occupational Mobility of Black Men in the United States, 1962–1973." *American Sociological Review* 49 (June): 308–22.

————. 1988. "More Universalism, Less Structural Mobility: The American Occupational Structure in the 1980s." *American Journal of Sociology* 93, no. 6 (May): 1358–1400.

Jencks, Christopher. 1990. "What Is the True Rate of Social Mobility?" Pp. 103–30 in *Social Mobility and Social Structure,* edited by Ronald L. Breiger. Cambridge: Cambridge University Press.

Klatzky, Sheila R., and Robert W. Hodge. 1971. "A Canonical Correlation Analysis of Occupational Mobility." *Journal of the American Statistical Association* 66, no. 333 (March): 16–22.

Lawley, D. N 1959. "Tests of Significance in Canonical Analysis." *Biometrika* 46:59–66

Siegel, Paul M. 1971. "Prestige in the American Occupational Structure." Doctoral dissertation. University of Chicago, Department of Sociology.

Stevens, Gillian, and David L. Featherman. 1981. "A Revised Socioeconomic Index of Occupational Status." *Social Science Research* 10 (December) 364–95.

The Vertical Scaling of Occupations: Some Cautionary Comments and Reflections[1]

David B. Grusky and Stephen E. Van Rompaey
Stanford University

The purpose of this commentary is to locate the symmetric scaling of intergenerational continuity (SSIC) within a detailed typology of vertical scales. In the course of doing so, we establish that SSIC cannot be seen as a "corrected" prestige or socioeconomic scale, nor is it a convincing proxy for the total assets and resources embedded in occupations. We argue that SSIC might be reinterpreted as the dominant dimension underlying patterns of intergenerational contact; however, some of the competing "network scales" tap a broader range of relational processes, and it would therefore be difficult to salvage a major role for SSIC within this tradition. We conclude by asking whether SSIC should be privileged relative to alternative scales that maximize other bivariate correlations.

The casual observer of American stratification research probably sees the occupational scaling tradition as a spent force (see Grusky and Takata [1991] for a relevant review). Although there is clearly *some* truth in this characterization, we must of course be careful not to overstate the case. Indeed, despite the evidence that the wider discipline is moving away from unidimensional scaling, there are surely no signs of intellectual stagnation among the small band of scholars who continue to be committed to revising, extending, and rethinking the Duncan (1961) socioeconomic

index (SEI) and its various spinoffs. The simple fact of the matter is that we can no longer treat the post-Duncan scaling literature as if it were a minor footnote in sociological history.

It is unfortunate, then, that Rytina (1992; hereafter simply "Rytina") makes no sustained effort to locate his own work within this broader research tradition. From a purely strategic point of view, there is much to recommend in Rytina's approach; it would be counterproductive to advertise a full range of competing scales when the objective is to persuade the audience to adopt a new one. However, if we wish to understand the vertical scale advanced by Rytina, it is clearly important to locate it in the context of contemporary empirical and theoretical work. The purpose of our article is to do just this. In the course of surveying the field, we shall establish that SSIC is not, in fact, a socioeconomic scale at all, but instead is closely related to a long tradition of relational scaling that has roots in the work of Weber ([1922] 1968), Warner (1949), and Laumann (1966). After reviewing some of the competing scales within this tradition, we will show that there is nothing that leads one to prefer SSIC as a generic measure of the stratification hierarchy. We will then conclude our commentary by asking whether SSIC might nonetheless serve in more limited fashion as the scale of choice for researchers who seek to model the process of occupational attainment.[2]

It should be emphasized that our commentary will *not* address the "empirical warrant" for SSIC. This warrant rests, so far as we can determine, on Rytina's claim (Secs. IV–VI) that the father-to-son correlation becomes stronger when SSIC is applied.[3] While data of this kind might eventually provide evidence on the "construct validity" of SSIC and its competitors (e.g., Cronbach 1960), we would argue that there is little point in proceeding to such matters when the nature of the underlying construct is left unspecified. It is, in fact, one of the curiosities of the scaling literature that the merits of particular measures can be hotly debated without any of the participants venturing to define the phenomena being measured. In some of these debates, the real sources of difference may well reside at the conceptual level, yet in typical American fashion the terms of debate are nonetheless ratcheted down to purely operational issues.

[2] It goes without saying that SSIC provides useful evidence on the contours of the American mobility regime However, rather than interpreting the canonical coefficients in such limited terms, Rytina claims that they reveal a more fundamental form of occupational stratification It is the latter claim which is the topic of our commentary.

[3] We do not wish to endorse this claim. The results presented by Hauser and Logan (1992) imply that the father-to-son correlation is strengthened under SSIC because it capitalizes on chance fluctuations in a relatively sparse mobility table.

We should note that the conceptual warrant for SSIC is hardly advanced by writing in vague generalities about the need to "correct" existing scales and thereby tap a more "fundamental" vertical hierarchy (e.g., Rytina, p. 1659). If we were to take this type of language seriously, it would seem to presuppose a *single* master hierarchy which underlies all systems of social action and behavior. Although a hierarchy of this kind might exist in some traditional societies, we think it is regressive to fall back on the same imagery in describing modern patterns of stratification.[4] This is not to say that one must err in the opposite direction by equating the concept underlying SSIC with its operationalization. There is little to be gained, for example, in claiming that SSIC "operationalizes the hypothesis that movement among positions in the division of labor is constrained by a hierarchy" (Rytina, p. 1668). What we need, instead, are some middle-range concepts that are neither so vague as to incorporate all possible vertical scales, nor so narrow that they become indistinguishable from the operationalization itself.

We shall proceed, therefore, by surveying some of the middle-range concepts that have been commonly deployed and by examining the relationship between these referents and the SSIC scale.[5] As we move through the discussion, it may be helpful to use figure 1 as a basic roadmap; the purpose of this figure is to identify some of the highlights in the history of occupational scaling and to locate SSIC in the context of these highlights. It should be self-evident that figure 1 was constructed for rather particularistic purposes and, consequently, no claim of comprehensiveness can be made.[6]

[4] We do not mean to downplay the degree of commonality that *can* be found across occupational scales of all kinds The only claim here is that such inconsistencies as do exist may represent more than errors in measurement or artifacts of method (cf. Featherman and Stevens 1982, p. 155) We think that this interpretation provides a plausible account of the data (see, e.g , Hodge 1981, p 404); however, we cannot offer any definitive evidence in its support, since it is always possible that the interscale correlations will gradually converge to unity as new and improved instruments are invented.

[5] We are not suggesting that SSIC must fit into existing traditions before it can be regarded as a useful contribution. However, we still think it is important to review these traditions carefully, if only because Rytina himself attempts to locate SSIC within them. We will conclude our commentary by entertaining interpretations of SSIC that fall outside the conventional scaling traditions that Rytina invokes.

[6] The flow charts in fig. 1 represent four major traditions in occupational scaling. In constructing this figure, we not only ignored some of the important events within each tradition, but we also suppressed the conceptual linkages that cut across these traditions.

PRESTIGE SCALES

The watershed events within the long-standing tradition of prestige scaling are now part of standard sociological lore, and we see little point in outlining a well-rehearsed history yet again. In this article, it should suffice to emphasize that prestige scales refer to the symbolic domain of stratification; the purpose of these scales is to represent "collective perceptions and beliefs" about the structure of occupational hierarchies (Hope 1982, p. 1012). We hope that this claim is uncontroversial. Although some scholars continue to participate in debates about "what prestige scales . . . actually scale" (Hauser and Featherman 1977, p. 5), there is nothing in such debates that would lead one to deny that prestige scores have a cognitive or symbolic referent. In almost all cases, the latter claim is taken as a given, and the real disagreements pertain to the *underlying sources* of the cognitive mapping that these scores reveal.[7] The "liberal stance" is that judgments about prestige and social standing are sensitive to honorific considerations (Hope 1982; also, see Siegel 1971; Turner 1958), whereas the opposing camp argues that socioeconomic factors play a dominant role in structuring occupational perceptions (e.g., Hauser and Featherman 1977, p. 7; Goldthorpe and Hope 1974). We should stress, of course, that advocates of the latter position are *not* simple reductionists; the prevailing view within this camp is that prestige scores are "substantively different from socioeconomic status" (Featherman and Hauser 1976, p. 405; also, see Nam and Terrie 1982, p. 33; see Featherman and Stevens 1982, p. 155).

If this reading of the literature is accepted, it makes little sense to treat SSIC as a corrected prestige scale. The latter terminology can only be justified under the heroic assumption that the cognitive dimension is a perfect mirror image of the vertical hierarchy governing occupational exchanges. We see no reason to entertain this form of radical materialism; after all, it appears "on the face of it" that SSIC and prestige scales tap different domains, and there is surely nothing in the results presented by

---

[7] The same point might be made with regard to the well-known debates about the behavioral consequences of prestige judgments (e.g., Goldthorpe and Hope 1974; Hope 1982). In fact, when Goldthorpe and Hope argue that prestige scales fail to capture patterns of "deference, acceptance, and derogation" (1974, p. 5), they are merely affirming the symbolic nature of the occupational rankings revealed by such scales. These rankings will no doubt be highly correlated with the vertical dimension underlying patterns of social behavior (e.g., assortative mating); however, we cannot expect the correspondence to be perfect, given that a host of mediating institutional forces serve to shape and distort the structure of interpersonal relations We would agree, then, with Goldthorpe and Hope (p. 11) that prestige scales do *not* measure prestige as it is conceived within the broader sociological tradition.

PRESTIGE SCALES   SOCIAL NETWORK
                                        SCALES

```
┌─────────────────┐              ┌─────────────────┐
│ Short-List Rankings│            │   Subjective    │
│ (Counts 1925; North│            │ Distance Scales │
│ & Hatt 1947; Hodge,│            │ (Wilkinson 1929;│
│ Siegel, & Rossi 1964)│          │  Martin 1963;   │
│                 │              │ Laumann 1966)   │
└─────────────────┘              └─────────────────┘

        ┌─────────────────┐        ┌─────────────────┐
        │ Detailed Rankings│       │    Objective    │
        │  (Siegel 1971;  │        │ Composite Scale │
        │ Goldthorpe & Hope│       │ (Laumann 1966)  │
        │      1974)      │        └─────────────────┘
        └─────────────────┘
```

┌───────────────────┐ ┌──────────┐ ┌──────────┐
│ Alternative Short- │ │Internationa│ │Female Scale│
│   List Methods   │ │   Scale   │ │ (e.g., Bose│   ┌──────────────┐
│ (e.g., Coxon & Jones│ │(Treiman 1977)│ │   1985)   │   │ Friends Scale│
│ 1978; Kraus, Schild &│ └──────────┘ └──────────┘   │(Stewart, Prandy,│ ┌──────────────┐
│   Hodge, 1978)   │   │& Blackburn 1980)│ │Intergenerational│
└───────────────────┘   ┌─────────────────┐   └──────────────┘ │ Mobility Scales│
                              │ Updated Detailed │   │ (Levine 1990; │
                              │ Rankings (Nakao │   │ Rytina 1992) │
                              │   & Treas 1990) │   └──────────────┘
                              └─────────────────┘

FIG. 1.—Recent highlights in the history of unidimensional occupational scaling.

Rytina that would motivate us to abandon our instincts here. This line of reasoning suggests, then, that the Siegel scores pertain to a *conceptually distinct* dimension of stratification; and, consequently, the associated father-to-son correlation is not merely an attenuated measure of a more fundamental form of association (see Jencks [1990] for a closely related argument; also, see Treiman 1977, p. 211). We should note that Rytina may well accept this conclusion. Although he starts off by arguing that there are "no obvious grounds for identifying observed differences [between occupational scales] with theoretically meaningful distinctions"

SOCIOECONOMIC
SCALES

SCALES OF ASSETS
AND REWARDS

```
Pure Averaging Scales
(Edwards 1943; U.S.
Bureau of the Census 1963;
Nam, Powers & Glick 1964)
```

```
Disaggregated
Occupational Assets
(e.g., U.S. Department of
Labor 1939; 1949; 1965;
1977; Temme 1975; Spenner
1980)
```

```
Prestige-Based
Socioeconomic
Scales (Duncan
1961)
```

```
Updated
Averaging Scales
(Nam et al. 1975;
Powers &
Holmberg 1978)
```

```
Disaggregated
Scales (Tyree &
Smith 1978;
Sewell 1985;
DeGraaf,
Ganzeboom, &
Kalmijn 1989)
```

```
Transformed and Updated
Prestige-Based Scales
(Sørensen 1979; Featherman
& Stevens 1982)
```

```
Mediating Scale
(Ganzeboom, DeGraaf,
& Treiman 1991)
```

```
Index of Job
Desirability
(Jencks, Perman,
& Rainwater
1988)
```

FIG. 1.—Continued

(Rytina, p. 1663), at the end of the paper he identifies the Siegel scale with "honorific standing" (p. 1680) and attempts to interpret its (imperfect) correlation with SSIC in substantive terms.

SOCIOECONOMIC SCALES

However, when we turn to the socioeconomic tradition of occupational scaling, it appears that disagreements of a more serious kind emerge. It is here, for example, that Rytina suggests that SEI is an "inspired"

approximation to SSIC, yet one which ultimately fails to fully capture the "hard facts driving rank reproduction" (p. 1680). In advancing this claim, Rytina goes on to suggest that SEI is "ad hoc" (p. 1678) and rests on "more or less arbitrary conventions" (p. 1659), whereas the mobility-based approach of SSIC is "conceptually superior" (p. 1677) and "in tune with important deeper themes" (p. 1684). These analytic arguments are coupled with a so-called empirical warrant for SSIC (Secs. IV–VI) that proceeds from the well-known result that socioeconomic scales cannot fully account for origin-by-destination exchanges (see, for related results, Blau and Duncan [1967, pp. 67–75]). By contrast, Rytina notes that SSIC was *constructed* to provide the best (one-dimensional) representation of occupational exchanges, and it therefore "correct[s] the widely employed SEI scaling for modest but significant defects in representing intergenerational continuity" (p. 1659).

These are strong claims, and they merit careful scrutiny. We would begin our conceptual work by emphasizing that the history of socioeconomic scaling is more complex than Rytina allows. Indeed, whereas Rytina insists that socioeconomic scales are ad hoc and fail to capture any stock sociological concepts (e.g., social class), other commentators have claimed a broadly defined Weberian pedigree for the socioeconomic tradition (e.g., Nam and Terrie 1982). There is, no doubt, a certain logic to casting Duncan as a neo-Weberian; after all, Weber is typically seen as defining classes in terms of the "social and economic life chances which people experience" (Nam and Terrie 1982, p. 29), and this suggests that some type of synthetic socioeconomic scale provides the necessary "basis for indicating social class" (Nam and Terrie 1982, p. 29; also, see Hodge [1981] for related commentary). At the same time, it has to be conceded that the distinction between social and economic resources is, at best, peripheral to the Weberian project; and, moreover, the categorical schemes that define Weberian social classes (Weber 1968, p. 305) are inconsistent with a purely gradational hierarchy. While there may be a "family resemblance" between the socioeconomic and Weberian traditions, we are unpersuaded by claims that the Duncan SEI or its counterparts serve to *directly* operationalize the Weberian definition of class.

The conceptual lineage for socioeconomic scales can be traced more convincingly to some of the post-Weberian efforts at occupational classification. For our current purposes, we need only note that scholars working within this tradition have typically characterized the occupational structure in terms of two "principles of hierarchization" (Bourdieu 1984, p. 120); the recurring distinction is between a simple economic hierarchy of occupations and an (imperfectly correlated) "sociocultural" hierarchy (see DeGraff, Ganzeboom, and Kalmijn 1989, p. 58). This distinction

was introduced by Edwards (1933, 1943) and popularized by his intellectual descendants (e.g., Nam, Powers, and Glick 1964); moreover, it also appears in the postwar research on occupational situs (e.g., Glenn and Alston 1968), and it reappears in modified form with the work of Bourdieu (1984) and his contemporaries (e.g., DeGraff et al. 1989). We would argue, then, that Duncan was invoking the *same type* of distinction, if only implicitly, when he claimed that SEI has "face validity . . . in terms of its constituent variables" (1961, p. 115). In seeking this form of validation, Duncan was suggesting that SEI is more than merely a convenient approximation to NORC prestige ratings, the Duncan SEI can also be seen as the "best" one-dimensional representation of an intrinsically two-dimensional occupational structure.[8] Whatever the original intentions of Duncan (1961) might have been, we should further note that subsequent scholars have completely eliminated the prestige interpretation by constructing updated socioeconomic scales with alternative (nonprestige) weighting algorithms (e.g., Nam et al. 1964; Ganzeboom, DeGraaf, and Treiman 1991). The latter scales must be interpreted entirely in terms of their constituent variables and the socioeconomic dimensions that these variables allegedly tap.

The foregoing comments suggest that socioeconomic scales are far from being "ad hoc" or "arbitrary." To be sure, the Duncan scale and its counterparts cannot be motivated with the standard references to classical theorists, yet this alone is hardly an appropriate basis for disqualifying them. We have argued, instead, that scales of this kind serve to glue together the two fundamental hierarchies that underlie modern occupational structures.[9] It is possible that Rytina regards all such synthetic scales as intrinsically ad hoc; however, this type of argument has historically been difficult to sustain, if only because a simple one-variable scale can always be reduced to a synthetic one by decomposing it into "more fundamental" parts (see Popper 1968; Hope 1982, pp. 1012–13). In any event, it is clearly inappropriate to regard SSIC as correcting SEI, since this terminology implies that the exchanges in a mobility table are perfectly socioeconomic in their structure. Although one could always *define* these exchanges as intrinsically socioeconomic, to do so clearly subverts

---

[8] This interpretation appears to be widely accepted among contemporary students of occupational scaling (see Hodge 1981, pp. 405–7). It should be kept in mind, however, that a somewhat broader definition of socioeconomic status can be found within earlier scaling traditions (e.g , Chapin 1933, p. 3; Sewell 1940).

[9] It is important to be precise here Although we are arguing that socioeconomic scales can be motivated with sociologically meaningful concepts, an element of "arbitrariness" will nonetheless emerge in the course of *operationalizing* these concepts. We would be hard-pressed, of course, to choose between competing measures of socioeconomic status (see fig. 1) on the basis of theoretical considerations alone.

the conventional definition that has been invoked by generations of re-searchers (e.g., Edwards 1943). This is not to deny the well-known result that SEI is highly correlated with the first vertical dimension underlying mobility exchanges (see Klatzky and Hodge 1971; Hauser and Feather-man 1977). What is unclear, however, is whether anything is to be gained by redefining the small departures from SEI as "fundamentally" socio-economic in nature.[10]

GLOBAL MEASURES OF POSITIONAL REWARDS

At some points in his text, Rytina suggests that SSIC should *not* be seen as a pure socioeconomic scale, since the "assets and advantages" that it captures are partly nonsocioeconomic in nature (see, esp., pp. 1680–81). This fallback interpretation appears to be a post hoc rationale for SSIC; in fact, if Rytina were truly interested in constructing a global measure of positional assets, we would have expected him to contrast SSIC against existing scales that share a similar rationale. As indicated in figure 1, the index of job desirability (IJD) is perhaps the most serious competitor to SSIC (see Jencks, Perman, and Rainwater 1988); the IJD measures global "success in the labor market" by summing the total utility of 14 mone-tary and nonmonetary job characteristics.[11] The inquiring reader will naturally want to know why SSIC should be preferred relative to compet-ing scales of this kind (see Temme [1975] for a relevant commentary).[12]

It should be clear, of course, that Rytina cannot claim that SSIC pro-vides a *direct* measurement of the assorted assets embedded in occupa-tions. Unlike the IJD, the SSIC taps a relatively limited domain of occu-pational life; the "row scores" generated under a standard canonical correlation are simply the expected mobility trajectories for children born into each of 308 origin occupations. In the absence of compelling evidence to the contrary, there is surely no reason to believe that these trajectories, in themselves, form the principal basis by which individuals judge the desirability of occupations. We have to agree with Parkin that "self-

[10] If SSIC were recalculated in a sufficiently large sample, we might find that the correlation with SEI approaches unity (see Hauser and Logan 1992). In this case, the distinction between the two scales becomes empirically trivial, but there would still be no conceptual basis for privileging the values that emerged under SSIC.

[11] The scale values for the IJD pertain to jobs rather than occupations (see Jencks et al. [1988] for details). It would be possible, however, to construct an occupational analogue to the IJD by pooling scale values across the jobs within each detailed occupation.

[12] We could always interpret SSIC as an index of the "specialized assets" that are salient for intergenerational mobility The issue at stake here is whether assets of this particular kind should be conferred a privileged status in building a general vertical scale.

reproduction . . . is not the overriding aim" of the modern worker (1979, p. 62); to be sure, most parents have a positive interest in transferring positional advantages to their children, yet the ability to do so probably ranks relatively low in the list of assets that they would consider when choosing between competing jobs (see Reiss [1961, pp. 29–31] for a relevant discussion).[13]

If a case for SSIC is to be made, it would therefore have to rest on the premise that patterns of interoccupational exchange provide *indirect* evidence on the "assets and advantages" embedded in occupations. This case makes for compelling reading (see Rytina, pp. 1680–81), yet we are ultimately unpersuaded by the (implicit) claim that mobility exchanges are governed by a pure vertical hierarchy. We suspect, in fact, that if this line of argumentation were made explicit, it would bear some similarity to a human capital model of market wage rates (e.g., Becker 1964). It appears that Rytina regards the mobility regime as an error-free arbiter of positional resources; that is, just as the ideal-typical economist expects wage rates to reflect the human capital "stored" in individuals, so too Rytina sees mobility rates as a realization of the social assets embedded in occupations. While models of this sort clearly provide important insights into the basic principles underlying patterns of exchange, they necessarily ignore the complex social forces that can distort the workings of the market. In the context of mobility exchanges, these forces are revealed in various "social bridges" that emerge between occupations that recruit from the same industrial, regional, or ethnic labor markets. The modern occupational structure is rife with bridges of this sort, and consequently there are "channels of mobility . . . that complicate the patterns of [occupational] movement as compared to what can be expected on the simple metaphor of a social elevator" (Blau and Duncan 1967, p. 117; also, see Goldthorpe and Hope 1974, p. 10). Although the hierarchy expressed in SSIC will be highly correlated with conventional occupational hierarchies (e.g., SEI), the observed discrepancies may well reflect idiosyncratic bridges rather than errors in the a priori representation of the vertical dimension.[14] We are well advised, then, to maintain an analytic distinction between the objective properties of occupations and the mobility experience of their incumbents (see Hauser and Logan 1992, pp. 1705–6).

[13] We are unconvinced, then, by the claim that SSIC captures the concept of "life chances" in some meaningful fashion (e.g., Weber 1968, p. 302). Although the row scores in a canonical correlation pertain to the life chances that are conferred upon children, we see no reason to assume that the life chances of *adults* are also indexed by the "reproductive potential" of their occupations.

[14] It would be overly optimistic to assume that the distorting effects of these affinities will only show up in the *secondary* variates of the canonical correlation

SOCIAL NETWORK SCALES

Up to now, our approach to SSIC and its analogues (e.g., Levine 1990) has been critical in tone, since our objective was to counter the claim that existing vertical scales are merely error-ridden approximations to mobility-based measures. Whatever the merits of our criticisms might be, we should nonetheless stress that our evaluation of SSIC is not an entirely negative one. It would appear, in fact, that Rytina has over-looked an interpretation for SSIC that has some sociological substance.[15] As indicated by the flow charts in figure 1, the approach advocated by Rytina falls within an *existing* scaling tradition; it turns out that SSIC is a relational scale indexing the "actually observed differential association of persons" (Laumann 1966, p. 90; also, see Lauman and Gutmann 1966; Mayer 1977).[16] If this simple behavioral interpretation is accepted, the SSIC can be defined as the (dominant) vertical dimension underlying patterns of intergenerational contact (see Wilkinson 1929; Martin 1963; Laumann 1966; Stewart, Prandy, and Blackburn 1980). Moreover, when groups of occupations cluster together on this dimension, we can speak of "social classes" in which intergenerational contact is easy and typical (see Weber 1968, p. 302; Laumann 1966). The classes so defined are merely groupings of detailed occupations held together by virtue of cross-cutting intergenerational networks.[17]

This approach to vertical scaling has its roots in the community studies of Warner (1949) and his contemporaries (e.g., Davis, Gardner, and Gardner 1941; also, see Warner 1963). However, when the discipline abandoned these early studies in favor of large-scale surveys, Laumann (1966) argued that similar types of scales could be constructed from the raw data in occupational cross-classifications (see, for a related scale, Stewart et al. [1980]). The important point for our purposes is that SSIC has the *same* structure as these modern-day relational scales. We are unpersuaded, therefore, by the claim that SSIC captures a vertical di-

[15] We will not be arguing that the *particular* scale values generated by Rytina can be safely interpreted. If we wished to apply SSIC in a serious fashion, we would clearly have to calculate new scores for a larger sample (see Hauser and Logan 1992).

[16] It could well be argued that relational scales come closer than prestige scales to tapping behavioral patterns of "deference, acceptance, and derogation" (Goldthorpe and Hope 1974, p. 5). In this regard, it is clear that prestige scales are vulnerable to criticism (see n 5 above), whereas socioeconomic scales are on somewhat safer conceptual ground

[17] The obvious irony here is that a scale of this kind cannot address the types of questions that have typically motivated research on occupational attainment. Indeed, insofar as the objective of such research is to determine "who gets the good jobs" (e.g., Jencks et al. 1988), we would surely prefer a vertical scale that captures the various monetary and nonmonetary assets that make jobs desirable.

mension (e.g., socioeconomic status) that is conceptually distinct from the social hierarchy underlying patterns of interoccupational contact (see, for a similar point, Stewart et al. [1980, p. 72]; also, see Laumann 1966, p. 29). Unless Rytina can specify the forces that would generate isomorphic hierarchies across different social arenas, we are well advised to adopt a direct interpretation of SSIC and its counterpart scales (e.g., see Levine 1990).

If this line of reasoning is accepted, the question that immediately arises is whether SSIC can successfully compete against existing scales within the Warner-Laumann tradition. We should note that SSIC is hardly an exhaustive measure of relational processes; after all, the exchanges in mobility tables reflect merely one form of interoccupational contact, whereas the original relational scales were defined "in terms of such social relationships as consanguineal and affinal kinship, friendship, informal social groups and cliques, and common [residence]" (Laumann 1966, p. 3). The SSIC might be seen, then, as the logical conclusion of a growing tendency among stratification researchers to assume that mobility processes define the most fundamental features of modern class systems. In recent decades, it has become increasingly acceptable to imbue mobility tables with all sorts of deep and wide-ranging theoretical implications, while other forms of stratificational data have been relegated to a purely secondary status (see Grusky and Fukumoto 1989, pp. 226–27; Breiger 1981; Levine 1990; also, see Goldthorpe 1987). The cynic might suggest that this tendency to overtheorize the mobility table only serves to legitimate our methodological interests in analyzing cross-classified arrays (see Hindess 1981; Johnson and Rattansi 1981). While the latter view is perhaps a bit extreme, we would certainly question the assumption that mobility tables index the most salient social networks. Unless we truly believe that parent-child ties are the only "social glue" holding classes together, there is no reason to give precedence to the relationships in a mobility table relative to those represented in tables of assortative mating, residential segregation, and the like (see Giddens [1973, pp. 107–12] for related comments).[18]

CONCLUSION

The purpose of our commentary has been to survey some of the middle-range referents for occupational scales and to examine their relationship

---

[18] As far as we can tell, the SSIC was not conceived by Rytina as a simple relational scale, nor was it ever interpreted in ways that would be consistent with such a conception. It would therefore be unfair to criticize SSIC for failing to "measure up" in this respect. The point of our comments is simply to make explicit the relationship between SSIC and competing occupational scales

to SSIC. In the course of doing so, we established that SSIC is not a socioeconomic scale at all, nor is it a convincing proxy for the assets and advantages that occupations confer. We also suggested that SSIC might be reinterpreted as a simple measure of interoccupational contact; however, it is unlikely that a measure of this kind will ever attract a broad following, if only because there are competing scales (e.g., Laumann 1966) that tap a wider range of relational processes. We should hasten to add that Rytina does not address middle-range issues of this kind in any meaningful way. The so-called analytic warrant for SSIC (Rytina, pp. 1677–80) is merely a restatement of the assumptions underlying this approach; we are told, for example, that the "ease and difficulty of intergenerational transition is distance made manifest," and that the "pattern of mobility is the empirical realization of a hierarchical arrangement of occupations" (Rytina, p. 1678). Although these statements testify to the wide variety of ways in which an assumption can be worded, they hardly qualify as meaningful argumentation on its behalf.

It may seem unfair to focus so relentlessly on analytic issues when Rytina promised an "empiricist critique of SEI" (p. 1660). It turns out, however, that Rytina never delivered on this promise; indeed, the so-called empirical warrant for SSIC is hopelessly confounded with analytic issues, since it proceeds by "correcting" SEI against a criterion variable that is *selected* in terms of pre-empirical considerations. In the absence of a well-defined rationale for this particular correction, we might reasonably prefer to recalibrate SEI against some alternative criterion (e.g., wealth, autonomy, authority). The obvious implication of these comments is that there are as many "empiricist critiques" of SEI as there are criterion variables against which SEI might be recalibrated. If Rytina wishes to single out one of these corrections, he needs to provide some analytic basis for doing so.

The case for SSIC appears to rest, at least in part, on the assumption that occupational scales should "directly operationalize ascription" (Rytina, p. 1679). As we have already noted, we find this to be a peculiar stricture to impose, given that the ability to transfer positional assets probably ranks relatively low on the list of things that workers value (see "Global Measures of Positional Rewards" above). Moreover, even if we *were* willing to accept this stricture, there is still no reason to privilege SSIC relative to alternative scales that maximize other forms of ascription. In the simple three-variable models estimated by Rytina (pp. 1672–76), there are two "ascriptive correlations" that might be maximized; the obvious competitor to SSIC is a canonical scale constructed from the correlation between occupational origins and schooling.[19] It would be

[19] This scale should be based on the canonical coefficients for the 308 occupational

straightforward, of course, to construct an empirical warrant for this alternative scale, since it necessarily corrects for nonlinearities in the effects of parental SEI on educational attainments. At the same time, the mediating role of education will be strengthened when these new values are applied, and the substantive conclusions favored by Rytina (pp. 1672–76) will therefore disappear. The moral to our story, then, is that sociologists should apply vertical scales that capture middle-range concepts of some sociological significance. If a researcher decides to rely entirely on an empiricist critique (Rytina, pp. 1659–60), then we are left with the impossible task of adjudicating between the various "corrections" that might be offered.

REFERENCES

Becker, Gary. 1964. *Human Capital*. New York. Columbia University Press.
Blau, Peter M., and Otis D. Duncan 1967. *The American Occupational Structure*. New York: Wiley.
Bose, Christine E. 1985. *Jobs and Gender: A Study of Occupational Prestige*. New York: Praeger.
Bourdieu, Pierre. 1984. *Distinction: A Social Critique of the Judgement of Taste*. Cambridge, Mass.: Harvard University Press.
Breiger, Ronald L. 1981. "The Social Class Structure of Occupational Mobility." *American Journal of Sociology* 87:578–611.
Chapin, F. S. 1933. *The Measurement of Social Status*. Minneapolis: University of Minnesota Press.
Counts, George S. 1925. "The Social Status of Occupations: A Problem in Vocational Guidance." *School Review* 33:16–27.
Coxon, Anthony P. M., and Charles L. Jones. 1978. *The Images of Occupational Prestige*. New York: St. Martin's.
Cronbach, Lee J. 1960. *Essentials of Psychological Testing*. New York: Harper & Row.
Davis, Allison, Burleigh B. Gardner, and Mary R. Gardner. 1941. *Deep South: A Social Anthropological Study of Caste and Class*. Chicago. University of Chicago Press.
DeGraaf, Paul M., Harry B. G Ganzeboom, and Matthijs Kalmijn. 1989. "Cultural and Economic Dimensions of Occupational Status." Pp. 53–74 in *Similar or Different?* edited by W Jansen, J. Dronkers, and K. Verrips. Amsterdam: SISWO.
Duncan, Otis D. 1961. "A Socioeconomic Scale for All Occupations." Pp. 109–38 in *Occupations and Social Status*, edited by Albert J. Reiss, Jr. New York: Free Press.
Edwards, Alba M. 1933. "A Social Economic Grouping of the Gainful Workers of the United States." *Journal of the American Statistical Association* 28:378.
———— 1943. *Comparative Occupation Statistics for the United States: 1870–1940*. Washington, D.C.: Government Printing Office.
Featherman, David L., and Robert M. Hauser. 1976. "Prestige or Socioeconomic

categories. If we are willing to condition on a linear scaling of EDUC (see Rytina, p. 1679), the occupation-by-education correlation can be maximized by fixing the scale values at the respective means (on EDUC) for each origin category.

Scales in the Study of Occupational Achievement?" *Sociological Methods and Research* 4:403–22.

Featherman, David L., and Gillian Stevens. 1982. "A Revised Socioeconomic Index of Occupational Status: Application in Analysis of Sex Differences in Attainment." Pp. 83–127 in *Measures of Socioeconomic Status,* edited by Mary G. Powers, Boulder, Colo.. Westview.

Ganzeboom, Harry B. G., Paul M. DeGraaf, and Donald J. Treiman. 1991. "A Standard International Socio-economic Index of Occupational Status." Unpublished paper. University of California, Los Angeles, Department of Sociology.

Giddens, Anthony. 1973. *The Class Structure of Advanced Societies.* New York: Harper & Row.

Glenn, N., and J. P. Alston. 1968. "Cultural Distance among Occupational Categories." *American Sociological Review* 33:365–82.

Goldthorpe, John H. 1987. *Social Mobility and Class Structure in Modern Britain.* Oxford: Clarendon.

Goldthorpe, John H., and Keith Hope. 1974. *The Social Grading of Occupations: A New Approach and Scale.* Oxford: Clarendon.

Grusky, David B., and Ann A. Takata. 1991. "Social Stratification." Pp 1955–70 in *The Encyclopedia of Sociology,* edited by Edgar F. Borgatta. New York: Macmillan

Grusky, David B., and Ivan K. Fukumoto. 1989. "A Sociological Approach to Historical Social Mobility: A Research Update." *Journal of Social History* 23·221–32.

Hauser, Robert M., and David L. Featherman. 1977. *The Process of Stratification.* New York· Academic.

Hauser, Robert M., and John A. Logan. 1992. "How Not to Measure Intergenerational Occupational Persistence." *American Journal of Sociology,* in this volume, pp. 1689–1711.

Hindess, Barry. 1981. "The Politics of Social Mobility." *Economy and Society* 10:184–202.

Hodge, Robert W. 1981 "The Measurement of Occupational Status." *Social Science Research* 10:396–415.

Hodge, Robert W., Paul Siegel, and Peter H. Rossi 1964 "Occupational Prestige in the United States, 1925–63." *American Journal of Sociology* 70:286–302.

Hope, Keith. 1982. "A Liberal Theory of Prestige." *American Journal of Sociology* 87:1011–31.

Jencks, Christopher. 1990. "What Is the True Rate of Social Mobility?" Pp. 103–30 in *Social Mobility and Social Structure,* edited by Ronald L. Breiger. Cambridge: Cambridge University Press.

Jencks, Christopher, Lauri Perman, and Lee Rainwater. 1988. "What Is a Good Job? A New Measure of Labor-Market Success." *American Journal of Sociology* 93:1322–57.

Johnson, Terry, and Alil Rattansi. 1981. "Social Mobility without Class." *Economy and Society* 10:203–34.

Klatzky, Sheila R., and Robert W. Hodge. 1971. "A Canonical Correlation Analysis of Occupational Mobility." *Journal of the American Statistical Association* 66:16–22.

Kraus, Vered, E. O. Schild, and Robert W. Hodge. 1978. "Occupational Prestige in the Collective Conscience." *Social Forces* 56:900–918.

Laumann, Edward O. 1966. *Prestige and Association in an Urban Community.* Indianapolis: Bobbs-Merrill.

Laumann, Edward O., and Louis Guttman. 1966 "The Relative Associational Contiguity of Occupations in an Urban Setting." *American Sociological Review* 31:169–78.

Levine, Joel H. 1990. "Measuring Occupational Stratification Using Log-Linear Distance Models." Pp. 208–24 in *Social Mobility and Social Structure*, edited by Ronald L. Breiger. Cambridge. Cambridge University Press.

Martin, John W 1963. "Social Distance and Social Stratification." *Sociology and Social Research* 47:179–86.

Mayer, Karl Ulrich. 1977. "Statushierarchie und Heiratsmarkt." Pp. 155–232 in *Klassenlagen und Sozialstruktur*, by Johann Handl, Karl Ulrich Mayer, and Walter Müller. Frankfurt: Campus.

Nakao, Keiko, and Judith Treas. 1990. "Computing 1989 Occupational Prestige Scores." General Social Survey Methodological Report no. 70. National Opinion Research Center, Chicago.

Nam, Charles B., John LaRocque, Mary G. Powers, and Joan J. Holmberg. 1975. "Occupational Status Scores: Stability and Change." *Proceedings of the American Statistical Association*, pp. 570–75.

Nam, Charles B., Mary G. Powers, and Paul C. Glick. 1964. "Socioeconomic Characteristics of the Population: 1960." *Current Population Report Series*. Washington, D.C.: Government Printing Office.

Nam, Charles B., and E. Walter Terrie. 1982. "Measurement of Socioeconomic Status from United States Census Data." Pp. 29–42 in *Measures of Socioeconomic Status*, edited by Mary G. Powers. Boulder, Colo.: Westview.

North, C. C , and P. K. Hatt. 1947. "Jobs and Occupation· A Popular Evaluation." *Opinion News* 9:3–13.

Parkin, Frank. 1979. *Marxism and Class Theory: A Bourgeois Critique*. New York. Columbia University Press.

Popper, Karl. 1968. *The Logic of Scientific Discovery*. London: Hutchinson.

Powers, Mary G., and Joan J. Holmberg. 1978. "Occupational Status Scores: Changes Introduced by the Inclusion of Women." *Demography* 15.183–204.

Reiss, Albert J., Jr. 1961. *Occupations and Social Status*. New York: Free Press.

Rytina, Steve. 1992. "Scaling the Intergenerational Continuity of Occupation· Is Occupational Inheritance Ascriptive after All?" *American Journal of Sociology*, in this volume, pp 1658–88.

Sewell, William H. 1940. "A Scale for the Measurement of Farm Family Socioeconomic Status." *Southwestern Social Science Quarterly* 21:125–37.

Sewell, William H., Jr. 1985. *Structure and Mobility: The Men and Women of Marseille, 1820–1870*. Cambridge: Cambridge University Press

Siegel, Paul M. 1971. *Prestige in the American Occupational Structure* Doctoral dissertation. University of Chicago, Department of Sociology.

Sørensen, Aage B. 1979. "A Model and a Metric for the Analysis of the Intragenerational Status Attainment Process." *American Journal of Sociology* 85:361–84.

Spenner, Kenneth I. 1980. "Occupational Characteristics and Classification Systems: New Uses of the Dictionary of Occupational Titles in Social Research." *Sociological Methods and Research* 9:239–64.

Stewart, A., K. Prandy, and R M. Blackburn. 1980. *Social Stratification and Occupations*. New York: Holmes & Meier.

Temme, Lloyd V. 1975. *Occupation: Meanings and Measures*. Washington, D.C.: Bureau of Social Science Research.

Treiman, Donald J. 1977 *Occupational Prestige in Comparative Perspective*. New York: Academic.

Turner, R. H. 1958. "Life Situation and Sub-culture " *British Journal of Sociology* 3 299–320.

Tyree, Andrea, and Billy G. Smith. 1978. "Occupational Hierarchy in the United States: 1789–1969." *Social Forces* 56:881–99.

U.S. Bureau of the Census. 1963. "Methodology and Scores of Socioeconomic Status." Working Paper no. 15 Washington, D.C.: Government Printing Office.

U.S. Department of Labor. 1939. *Dictionary of Occupational Titles,* 1st ed. Washington, D C.: Government Printing Office.

———. 1949. *Dictionary of Occupational Titles,* 2d ed., vol. 1. Washington, D.C.: Government Printing Office.

———. 1965. *Dictionary of Occupational Titles,* 3d ed., vol. 1. Washington, D.C. Government Printing Office.

———. 1977. *Dictionary of Occupational Titles,* 4th ed. Washington, D.C.: Government Printing Office.

Warner, W Lloyd. 1949. *Social Class in America: A Manual of Procedure for the Measurement of Social Status.* New York: Harper & Row.

———. 1963. *Yankee City.* New Haven, Conn.: Yale University Press.

Weber, Max. (1922) 1968 *Economy and Society.* Berkeley. University of California Press.

Wilkinson, Forrest. 1929. "Social Distance between Occupations." *Sociology and Social Research* 13:234–44.

Response to Hauser and Logan and Grusky and Van Rompaey[1]

Steve Rytina
McGill University

I. SOCIAL FACT NOT ARTIFACT: REPLICATION
WITH SSIC IN OCGII

Hauser and Logan's (hereafter H&L) response to SSIC has many facets, but its core is a statistical critique. Reduced to essentials, their claims are that SSIC lacks cross-sample validity and that the apparent increase in the intergenerational correlation is due to over-response to sampling error. But their attempt to demonstrate this is flawed.

Symmetric scaling of intergenerational continuity is a procedure (or algorithm) for determining a hierarchical pattern from mobility data. It results in, as I state above, "SSIC scores only for the data on which they are based, and other more or less similar arrays would arise from different samples." This means that comparison across samples requires new estimates. But H&L omitted this central step and thus failed to correctly grasp cross-sample performance.

Direct estimation reveals four central contrary results. First, SSIC is empirically indistinguishable from application of a maximum-likelihood method due to Goodman (1979a, 1979b, 1981a, 1981b), and thus falls within a central tradition of mobility analysis. Second, application of SSIC to the OCGII datá reveals almost identical summary estimates, which is hardly consistent with an absence of cross-sample validity.[2]

---

[2] Here, "GSS" refers to the sample (derived from Davis and Smith [1986]) used in the original paper, while "OCGII" refers to an extract I prepared from the public domain version of the Occupational Changes in a Generation II data set provided by the Data Program and Library Service (1983) at Madison. It is restricted to respondents aged 25–64 in the experienced civilian labor force and to the 307 occupations that overlap with the original NORC data set. The results are weighted in a manner that reproduces results reported in Featherman and Hauser (1978). For the record, it is slightly different than that employed by H&L, who based their work on an unreleased version that includes some slight modifications. I would like to record here my

Third, direct estimation of test-retest reliability shows that the scale results are quite similar across the original and the replicated analysis. Fourth, it can be shown directly that the similarity of estimates across samples vastly exceeds what could be observed if measurement error was the source of the original findings.

This vitiates H&L's interpretation. Their central claim is lack of cross-sample validity arising from incorporation of sampling error. From this one could infer the futility of attempting replication. Yet replication does occur. From this one must infer the futility of their analysis strategy.

In due course, it will be seen that their interpretation rests on an incomplete examination of results across samples that founders when the fuller set of facts is in view. But the main focus of what follows is the more positive agenda of displaying evidence that SSIC leads to reliable results within and between samples. H&L's attack provides occasion to greatly extend the evidence motivating further investigation with SSIC.

Stability of Results across Differences in Occupational Frequency

A great advantage of SSIC compared with fixed scales like SEI is that fresh data may provide corrected estimates of occupational location with respect to mobility. This allows for the possibility of change between samples or over time. But such sample-based estimates are subject to variability.

The SSIC operates at two levels. The correlation estimated as the ultimate measure of intergenerational continuity rests upon estimates of occupational location. These, in turn, can be regarded as derived from the (overlapping) subsamples for each of the detailed occupations. These range in size from the 1,422 observations of farm owners and tenants to the six occupations represented by exactly one father and one offspring. Of course, estimates based on as few as two observations are fuzzy. But varying subsample frequency provides an indirect control over the contribution of such "error."

H&L's thesis is that over-response to particularities of the sample, or "overfitting," inflates the overall correlation. While such "error" is not directly observable, this hypothetical inflation should be more prevalent as one estimates more location parameters and/or bases estimates on smaller subsamples. Conversely, restricting analysis to fewer, better represented occupations should reduce contamination by error. On H&L's thesis, this should result in reduced correlation.

TABLE 1

INTERGENERATIONAL CORRELATIONS AFTER DELETION OF OCCUPATIONS THAT DO
NOT MEET MINIMUMS FOR MARGINAL FREQUENCY

Minimum Marginal Frequency	Fraction of Sample	Fraction of Occupations	SEI Correlation	SSIC Correlation
1	1 000	1 000 (308)	3382	.4505
2	8673	7760 (239)	3549	.4579
3	.8267	.6526 (201)	.3523	.4509
4	.7566	.5487 (169)	.3566	4521
5	.6991	.4643 (143)	.3609	4484
6	.6475	.4058 (125)	3584	4482
7	.6187	3701 (114)	3589	.4495
91occs	.5680	2955 (91)	3559	.4511

NOTE —This table is based on 7,965 cases from the NORC General Social Survey (GSS), 1972–86.
Set labeled "91occs" are those with at least 25 observations and a minimum marginal frequency of 6
Numbers in parentheses are N's for occupations

Table 1 shows otherwise. The table reports separate analyses for the mobility subtables defined by restriction of the original GSS data to occupations that meet criteria of minimal marginal frequency.[3] The overall marginal distribution is quite uneven, so that the more restricted analyses are based on much larger average numbers of observations per occupation. But extension to rare occupations does not inflate the SSIC correlation, which is essentially constant. The SSIC correlation actually varies less (SD = .003) than does that for SEI (SD = .007).

Such results run counter both to intuition and to H&L's central thesis. The next section provides some hint to a plausible basis for this observed robustness under extension to occupations where the sample provides very limited information.

Statistical Properties of SSIC

H&L unfavorably compare SSIC to association models for mobility tables primarily originated by Goodman (see H&L's n. 3 above). Of particular

---

[3] The last line of table 1 reports on a subsample of 91 occupations. With the further deletion of one occupation at the inclusion minimum, this subtable was small enough to fit within the 64K limit imposed by the computer technology on which the original analysis was carried out in 1988 The inclusion criteria was a symmetric minimum of at least six fathers and at least six offspring as well as at least 25 observations. In the following reports, analyses based on these occupations will be identified by the label "(91occs)"—for 91 occupations—to contrast with the "(308occs)" and "(307occs)" of the fuller analyses.

interest is Model II*, which is among the models first brought to sociological attention in Goodman's (1979a) companion to Duncan (1979). Model II* is Goodman's designation for a maximum-likelihood procedure for log-multiplicative modeling of association in a contingency table using symmetric locations for rows and columns. Goodman (1981b) showed the close affinity of the asymmetric analog, Model II, with bivariate normality.

When Model II* is applied to the GSS (308occs), the resulting estimates of location with respect to the pattern of association in the mobility table correlates .998 with SSIC. For all practical purposes, SSIC is the same as this maximum-likelihood procedure. This is detailed elsewhere (see Rytina 1991). Nearly the same convergence occurs in the OCGII, where the correlation between the results of the two procedures is .993.

Maximum likelihood, broadly speaking, makes optimal use of sample information. This provides a hint to why incorporation of fuzzy estimates for rare occupations has limited impact on the overall result. And it suggests that SSIC can be regarded as (nearly equal to) an optimal summary of the pattern in the raw material. Since that material is a mobility table, it seems quite reasonable to assert without qualification that the result describes "occupational mobility." H&L implicitly suggest that this analytic tradition, which they have on other occasions advanced, is on this occasion invalid.[4] I do not agree.

Cross-Sample Validation

A central oddity of the H&L critique is that "cross-sample validity" is denied without direct examination. The usual way to assess a procedure in two samples is to carry it out in both samples and compare results. The SSIC passes this conventional test of cross-sample validation to a very high standard.

OCGII (307occs) results in an intergenerational correlation of .45003, which is nearly identical to the .45048 found in the GSS. The narrower

---

[4] The exact details seem to reflect a gap between us in tacit assumptions. On page 1707 H&L state that "similarly, Hout's (1984) SAT model of the mobility table might be compared with log-linear or log-multiplicative models that generate intrinsic scalings of occupation." Since SSIC leads to essentially the same estimates, this passage implicitly advocates my analysis. But since they were unaware of the convergence, the intent of their prose is to indicate motivation for analyses they regarded as hypothetical (and preferred?) alternatives. Their continuation relegates such analysis to some inferior status as mere adjunct. Would this mean that empirical findings of different occupational location and so forth should be ignored? This seems to me unwarranted, but it is what a literal reading of their discussion suggests.

OCGII (91occs) results in .4490 while the GSS (91occs) yielded .4511. "Point four five" adequately summarizes the result. There is no meaningful difference between the samples.

A more direct measure of cross-sample performance is the similarity of scale results calculated independently in the several samples. As one indication of magnitudes, the correlation is .941 between Duncan's (1961) original SEI scores and the Stevens and Featherman's (1981) update called MSEI3.[5] The correlation is .8624 between two different measures of education found in the OCGII data.[6] The SSIC compares favorably. For the fuller set of occupations the test-retest correlation is .897. Restriction to the more frequent (and accurately estimable) 91 occupations results in a correlation of .931.

There is no absolute standard available, but the degree of cross-sample reliability is certainly respectable. As the example of the two SEI variants shows, it is commonplace within the fine print that "invariant" quantities come out somewhat different with different samples. As the example of education shows, the degree of inaccuracy for SSIC is considerably less than is ignored as a matter of course.

Figures 1 and 2 report various three-variable path models using different occupational scales and different measures of education based on OCGII. (The same results for GSS [308occs] are in my first article in this volume). Examination of these (and the implicit correlations) reveals that key findings generalize across the two samples. The SSIC consistently results in an increase in the "inheritance" of educational attainment as compared with SEI. Offspring SSIC is consistently less correlated with respondent education. Combined with the increase in the occupational continuity correlation, SSIC results in a greater weight on the direct, or ascriptive, effect.

The path diagram based on the OCG questionnaire education measure results in coefficients within 5% of the comparable GSS estimates. Even if space allowed, it would not be informative to defend this in detail, but it serves as striking illustration of the close similarities implicit in the replicate results not examined by H&L. This remarkable "coincidence" sets the stage for seeing how further parallels across samples left unexam-

---

[5] Correlations for occupation in this section are based on OCGII offspring frequencies. I would like to thank Robert Hauser for providing MSEI3 scores.

[6] What I call OCGeduc is from a question in the OCG questionnaire, and the wording is most similar to the GSS measure. This measure, however, has 1126 (unweighted) cases of missing data. It appears that most work based on this sample actually uses the measure from the CPS survey, a measure I label CPSeduc. This variable has no missing data in the subsample at issue and apparently includes allocated values.
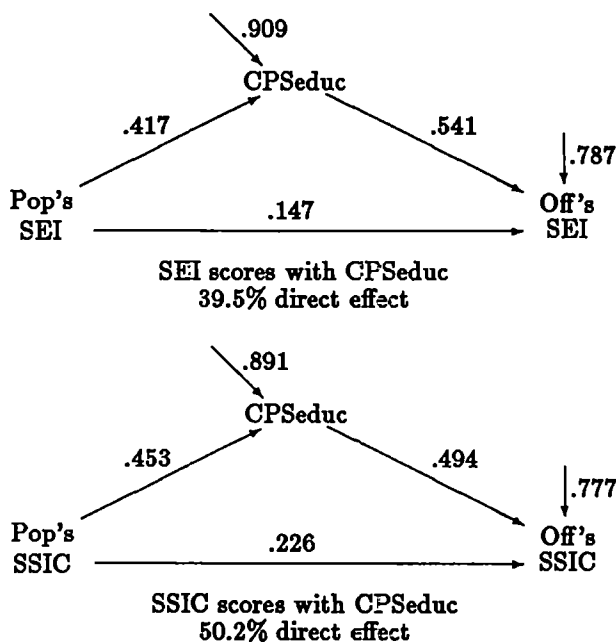
FIG. 1.—Path diagrams of the role of education in intergenerational continuity of occupational rank that are based on 23,138 raw cases weighted to 17,187 cases in the 307 × 307 mobility table extracted from the OCGII data.

ined in their critique shed light on the limited quantitative implications of unreliability due to sampling error.

Setting Bounds on the Implications of Unreliability due to Sampling Error

Allegations of bias due to sampling error are central to H&L's critique of SSIC, but the basis for such claims is indirect and incomplete. Table 2 reports relevant correlation estimates. The basis for their claims have been supplemented by parallel quantities for both OCGII and GSS.[7]

Parallel estimation reveals severe logical difficulties. Many of their claims supposedly rest on the properties of estimators and not just specific

[7] Because of the space limits of the journal, I am restricting my attention to the comparison of the OCGII and GSS data. That limited focus could be independently justified: any results that replicate across these two data sets suffice to show that the findings cannot be laid to method bias that capitalizes on GSS sampling error. But a longer discussion that touches upon other issues and data addressed by H&L is available from the author on request.

.916
OCGeduc
.400 .502
                                        |.806
Pop's   .173   Off's
SEI   ─────────────────────────▶   SEI

SEI scores with OCGeduc
46.2% direct effect

.901
OCGeduc
.433 .453
                                        |.795
Pop's   .253   Off's
SSIC ─────────────────────────▶   SSIC

SSIC scores with OCGeduc
56.3% direct effect
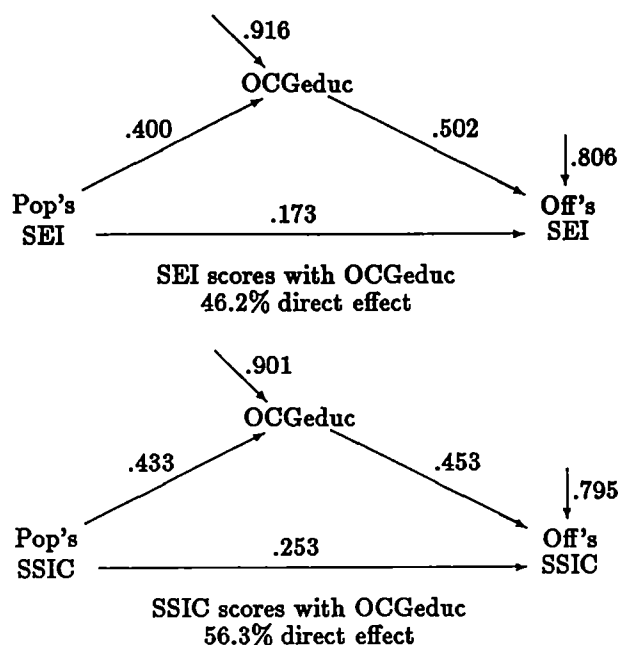
FIG. 2.—Path diagrams of the role of education in intergenerational continuity of occupational rank that are based on 22,002 raw cases weighted to 16,344 cases in the 307 × 307 mobility table formed from OCGII data that are not missing on the OCG education measure.

TABLE 2

Assorted Estimates of Intergenerational Continuity of Occupation Based on General Social Survey, 1972–86, and on Occupational Changes in a Generation II

	SAMPLE	
SCALE (SOURCE)	GSS	OCGII
SEI	.3382	.3727
CC	.4893	.4758
Lawley	.3376	4446
SSIC (GSS)	.4505	.3756
SSIC (OCGII)	3741	.4500

NOTE.—The GSS results are based on 7,965 cases used in the original analysis The OCGII results are based on 17,186.7 weighted cases found in occupations used in the GSS analysis. Coefficients not reported by H&L are underlined Estimates labeled "CC" are canonical correlations, while those labeled "Lawley" are deflated canonical correlations using a formula derived from Lawley (1959)

values. Where examination of both samples in parallel reveals that results depend on the particularities of single samples, the proposed generalizations are false. As an economy in presentation, references to parallel quantities omitted in their analysis will be underlined.

Two elements may be singled out to illustrate the difficulties with their argument. First, H&L cite Lawley (1959) for a downward correction to (the principal) canonical correlation.[8] The SSIC yields a (much) higher value in GSS. H&L assert that this excess must be regarded as "response to sampling error" or upward bias. Second, they calculate correlations using SSIC values from the GSS in fresh data, compare these with SEI correlations, and conclude that no real increase in correlation can be found.

In OCGII, the *SSIC correlation of .45* and the *Lawley correlation of .4446* are essentially identical. The "overfitting" that H&L assert as a property of SSIC fails to appear. This casts doubt on their claim that increased correlation observed with SSIC is artifact of method, since such would appear with all samples. For that matter, by their chosen criterion, the increase in correlation of SSIC over SEI is correct in magnitude.[9]

H&L's interpretation of the Lawley (1959) estimate is questionable in another regard. Although it is based on the distinct procedure of canonical correlation, they took the liberty of applying the value to SSIC. One must infer that they believe it yields a result that applies to any and all scaling procedures, in technical terms, an upper bound. The uncorrected canonical estimate on which it is based has this property.

If this property held, no numerical assignment of occupations that was both free of sampling error and exceeded this upper bound would be logically possible. Yet such a counterexample is apparent in table 2. The SSIC values based on OCGII are logically free of GSS sample error but yield a *correlation in the GSS of .374* that is well above the supposed upper bound of .338.[10]

---

[8] H&L persistently conflate canonical correlation and SSIC, although the two are distinct procedures that produce different results. Careful reading reveals that SSIC consistently outperforms canonical correlation in the respects of concern to them. Yet such statistical theory as they present is applicable only to canonical correlation. Such estimates of location as they carry out are canonical scaling. At best, this conflation of different methods means their critique is unfocused

[9] The estimates based on Lawley further "prove" that the correlation fell enormously from OCGII to GSS. One cannot simultaneously affirm the Lawley estimation strategy and the approximate invariance (and cross-sample validity?) of status attainment results. As I am not an advocate of the Lawley estimator, which rests on simplifying assumptions I find dubious for these data, I happily leave this problem to those who think it provides a sound tool for cross-sample evaluation.

[10] A downward correction to the result initially employed by H&L led to a Lawley value such that the SEI correlation is slightly (16%) larger and is another counterex-

The second central element of H&L's interpretation concerns such cross-correlations. In both data sets, SSIC from the opposite number results in a correlation of (about) .375. According to H&L, the similarity of the cross-correlation in OCGII to that of SEI demonstrated that no real increase was involved. Symmetrically, the *cross-correlation in the GSS* reveals that at least a modest increase occurs.

But H&L's interpretation of cross-correlations and within-sample correlations is flawed by a logical asymmetry. They assert, for example, that the increase in the GSS from a Lawley estimate of .338 to the .450 of SSIC is due to incorporation of the details of the GSS sample. This would mean that measured SSIC is 75.1% (.338/.450) "true" variance and 24.9% "error." Under this interpretation, SSIC based on GSS is deemed unreliable in definite degree.

This implies that the cross-correlation of .3756 reported by H&L for OCGII is based on an unreliable measure. Such measurement error necessarily attenuates empirical correlations. The implied correction is to divide by the .751 "reliability" implicit in the core, unflattering characterization of SSIC's apparent success in uncovering correlation in GSS. If H&L's interpretation is carried to its conclusion, the "true" value of the correlation must be .5001 (the ratio of .3756/.751), which is hardly consistent with the thrust of their argument.

The logic follows the maxim that one cannot have it both ways. The claim that .375 is a sound estimate, and not attenuated, implies that reliability is 100%, which means that .45 cannot be inflated by measurement error. Claims to the effect that .45 is too high because of unreliability implies a commensurate increase for .375. There is no logically sound way to interpret small cross-correlations as both evidence of unreliability *and* as a failure to detect increased occupational continuity. However, this unsound interpretation and the nonreplicable result using Lawley estimates are the heart of H&L's statistical critique.

At this logical dilemma, H&L's interpretation may be left aside for a more reasonable attempt to summarize cross-sample performance. It is feasible to derive point estimates of a *lower bound* for the correlation. Within each sample, the two marginal frequencies lead to two correlations between the local and externally estimated SSIC score. The geometric mean of these is the relevant reliability coefficient (Ghiselli, Campbell, and Zedeck 1981, p. 290).

These only provide a lower bound because the correction rests on the (quite unwarranted) assumption that all differences between samples are

---

ample. I will only note but not pursue the dubious implication that SEI is slightly beyond perfection. My text retains the original emphasis on the more dramatic illustration

"error." This assumes away real change and further attributes all non-sampling differences (which Jencks et al. [1979] show consistently far exceed chance expectations) to measurement error or, in this context, SSIC.

The results display further strong cross-sample parallelism. The reliabilities are .894 (GSS) and .904 (OCGII), and these yield corresponding corrected cross-correlations of .418 and .415. Results are more impressive for the more stable 91 occupations. In OCGII, the cross-correlation is .4051, the cross-scale correlation is .9405, and the corrected cross-correlation is .4307, which is quite close to the simple SSIC correlation of .4490. With the GSS, the cross-correlation is .4036, the cross-scale correlation is .9283, and the corrected cross-correlation is .4347, which is also quite close to the simple SSIC correlation of .4511.

The difference between samples is negligible. The remaining contrast is between the fuller and narrower sets of occupations. For both sets, the maximum-likelihood point estimate is .45. The difference is that the lower bound is lower, and hence uncertainty greater, for the set including rare occupations. For the narrower set, the same point estimate is accompanied by a lower bound that is negligibly smaller.

Recall, however, that the 91 occupations are a literal subtable. These results show that with minimal uncertainty one can demonstrate greater occupational continuity for this subset of empirically common occupations. Extension to the full set entails a larger interval of fuzziness, but the point estimate is unchanged. In both samples, for both frequent occupations and rare, there is considerable increase over SEI. Direct estimation, with measurement instability taken into account, sustains the central empirical conclusion that traditional scales understate occupational continuity.

An important counterpart is that SSIC captures a cross-sample replicated property of occupations that is not SEI. Table 3 reports path coefficients employing earlier measures to predict later ones. The first two lines show that updated MSEI3 values primarily reflect the parallel SEI results, but that the net impact of SSIC is negligible. The next two lines reveal nearly identical estimates for SSIC, except that roles are reversed. Each construct lines up with its replicate but shows minimal net impact of the alternative.[11]

H&L's thesis would imply that the GSS-SSIC results are an error-ridden variant of SEI. But the net effect for SSIC across samples is calculated from the residuals with respect to SEI for SSIC in each of the

[11] This pattern appears even though GSS-SSIC is highly correlated (.830 [308occs] and .863 [91occs]) with SEI. Thus the SSIC results are both closely parallel to SEI and clearly distinct.

TABLE 3

PATH COEFFICIENTS DESCRIBING THE SIMILARITY AMONG VARIOUS OCCUPATIONAL
RANKINGS FROM CORRELATIONS BASED ON OCGII OCCUPATIONAL FREQUENCIES

	INDEPENDENT VARIABLES	
N OF OCCUPATIONS AND DEPENDENT VARIABLE	SEI	SSIC(OCGII)
307 occ·		
MSEI3 ...	.782	177
SSIC (GSS)	.153	754
91 occ.		
MSEI3	871	091
SSIC (GSS)	053	.882

samples. On the H&L thesis, this net effect is a (part) correlation between
two sets of logically independent errors from separate samples and should
be about zero. This expectation is overwhelmingly disconfirmed. The net
effect of OCGII-SSIC is massive, while that of SEI is negligible. This
shows that the GSS locations are both different from SEI and reflect
properties of occupations evident in the OCGII. Of course, the estimated
OCGII locations are pristine with respect to GSS sampling error. Thus
the revised pattern of locations revealed by SSIC cannot be explained
away as an artifact of the GSS sample.

On the Watershed Assumption of Fixed Occupational Scales

The key to the preceding analysis is to regard SSIC results as a function
of data and then to examine how estimates vary when additional data is
considered. H&L did not undertake this. Implicitly, they drew on an
assumption that has been central to status attainment research: occupa-
tional rank is a firmly established fixed scale.

Five further examples may be cited which together illustrate the cen-
trality of the assumption for H&L. First, they take the GSS (1972–86)
results as generic and apply them in other samples. Yet OCGII is much
the larger data set and therefore the preferred source if one insists on a
single answer. But, as just indicated, the central steps in their argument
collapse when the role of source and replicate are reversed.

Second, even while they insist, with some good reason, that as sample
statistics SSIC values must be subject to some measure of variability,
they did not integrate such unreliability in interpreting their central de-
vice of cross-correlations. Their case against SSIC requires that the val-
ues from one relatively small sample, and some subsamples as tiny as
two cases, can be treated as fixed values.

1739

Third, they did not address changes in values implicit in replications or updatings of SEI. Such changes are indeed somewhat less than with SSIC. But the SSIC results are quite nearly as stable for frequent occupations. It is not the case that one measuring rod varies while the other is rigid.

Fourth, while they appear to advocate log-multiplicative approaches that effectively rescale occupational locations, they label such as "adjuncts" (cf. n. 3 above). It is unclear what this precisely means, but it seems to suggest that no empirical modification of accepted scale values should be entertained.

The fifth example is perhaps the subtlest. H&L state (p. 1706) that an occupation upgraded because of advantaged origins would be a locus of greater downward mobility. An example to which this should apply is that physicians were raised on SSIC on such grounds.

Physicians are raised because their origins (and destinations) are more advantaged than is consistent with their SEI placement. Allowing for the complications of regression to the mean, H&L's assertion is quite consistent if one assumes that physicians occupy a known level, which implicitly is the SEI value. Then offspring who become physicians are effectively using up greater origin advantages to obtain nothing more than the established level and thus are "downwardly mobile." If this sounds contrary to intuition, a probable source is that the reader, like me and unlike H&L, is willing to allow that physicians might occupy a rank somewhat different than that assigned by SEI.

By addressing rank as a variable SSIC simultaneously meets three central rationales within the established positivist frame of reference. It is (essentially) maximum likelihood. It boosts explained variance to a maximum, exceeding extant alternatives.[12] It replicates across samples.

These attractive properties are readily demonstrated but only through relaxing the fixed scale assumption. The SSIC values from a single sample are not population values. H&L failed to accommodate this. They did not reestimate nor consistently explore the implications of measurement instability (across samples, unreliability plus true change). What proves invalid, then, is their assumption that SSIC results from one sample might be freely transported to another as if they were population values. What proves valid, that is, replicable, are the paired assertions that

[12] An indication of the centrality of this assumption is Jencks's (1990) succinct summary of Featherman, Jones, and Hauser (1975) and Featherman and Hauser (1976). "They define an occupation's true rank as whatever best explains the correlations between the educational attainments and occupations of fathers and sons" (Jencks 1990, p. 109). Since SSIC is superior with respect to occupation and does slightly better at "explaining" offsprings' educational outcomes, one could make the case that it meets these standards.

traditional scales underestimate occupational continuity and that locations with respect to mobility chances are inadequately summarized by SEI.

Estimates from SSIC are not literally invariant across samples used for estimation. In principle, SEI is equally a set of estimates and does change with new samples. But in practice, SEI is based on huge census data sets and is decenially invariant. Employment of SSIC does entail some loss of comparability across studies, though the certainty lost is partly an illusion of convention. In my view, this is a price worth paying to have a full accounting of intergenerational continuity. The instability in estimation is certainly not worse than for many measures in common use, and probably better than most. An extra degree of freedom (or 308, if you will) across research reports is not a price out of line with the advantages.

On the Meaning and Interpretation of Scales

H&L regard SEI as distinctively established as a real, or tangible, attribute of occupational titles. Our disagreement here is slight. This complex compound of (approximately) average education and income of occupational incumbents is very solid, and certainly it is one of the outstanding accomplishments of sociological measurement. But, as Hodge (1981) pointed out, it is ad hoc. The weakness is the absence of a compelling rationale for confining attention to education and income among the many attributes of occupation.

The SSIC is essentially just a different sort of average, namely of occupational origins and destinations. In this case, the rationale is that distance between origin and destination when both are construed as locations in a social hierachy is one-to-one with the concept of occupational mobility. It is unclear why H&L regard this as "deliberately disregard[ing] the existential, measurable aspects of life experienced in occupations" (p. 1706), unless they somehow regard the social origins of incumbents as an external mishap with no bearing on work assignments and rewards. As I pointed out in the original paper, the final result is quite similar to simple averages over origins and destinations scaled by SEI (or prestige for that matter), but is ultimately freed from dependence on any particular prior notion of occupational location by climbing the gradient of likelihood toward the maximum.

Does this mean that an occupation that scores higher on SSIC is genuinely more advantaged? Symmetry comes into play here.[13] Higher SSIC

[13] Rytina (1991) reports tabular analyses that empirically sustain the assumption of symmetry.

(higher than, e.g., SEI) means that offspring do better. This implies the existence, for the *average descendent,* of assets with respect to mobility. Such an average result is, by common definition, "ascribed." Further-more, higher SSIC means that the occupation recruits more successfully from strata higher on SEI (or prestige, in empirical fact). This implies that the occupation is more attractive to those who possess the greater measure of assets in the competition for advantage. I see no great problem with the inference from such average choice to an underlying reality of advantage, perceived or even real.

The SSIC is further founded on the a priori theoretical notion that occupations are stratified into layers insofar as continuity occurs within families. Any coding of occupation into nominal categories delimits the unordered differentiation of the division of labor. The fundamental socio-biological relation of descent defines possible ascriptive assignment through kinship. The SSIC combines these to estimate the hierarchical arrangement through which succeeding generations repopulate the divi-sion of paid labor.

The clarity of the concept is most apparent with respect to the equality of occupations. Two occupations are equated insofar as they are equal in (average) origin and descent and thus equivalent as way stations in the intergenerational reproduction of the division of labor. With allow-ance for regression to the mean, equivalence also entails maximum co-access.

Equality on SEI is far less crisp. At an extreme, two occupations would be equated if one employed only persons with above-median education but paid them below-median income while the other provided universal high economic returns to persons with inferior educational backgrounds. Less extreme variants on this theme of compensating differences are ubiq-uitous.

Thus an occupation offering lesser economic return for more education is deemed equivalent to one that garners greater rewards for the less educated.[14] Up close, it is inescapable that this venerable standard in-volves an addition of apples and oranges.[15]

[14] A questionable implication is that if some type of persons was segregated into a subset of occupations and, in addition, got less income returns for education, the logic of SEI would transmute such disadvantage into a species of equality.

[15] Jencks (1990) drew on a more theoretically coherent index of job desirability (IJD) that employed popular perceptions to quantify trade-offs. One does, after all, add (and subtract) apples with oranges in filling an ideal market basket Subject to his necessity of employing a strong assumption of consensual evaluation, I would agree that his results show that the subjective advantage of jobs apparently is less rigid than occupational assignment, which is a result that parallels, partly by construction, the well-known lesser rigidity of measured economic status.

The SEI nevertheless has great pragmatic value. But it remains a constructive cousin of prestige, whose most prominent virtue was greater explanatory power in mobility studies. The SSIC is not conceptually identical to either prestige or SEI but, as shown in table 3, is at once very similar to but distinct from SEI. And it does the central task better. It is precisely in the context of mobility studies that it "corrects SEI" by providing a more adequate description of the issue at hand, namely, By what pattern is occupation intergenerationally reproduced?

Is Occupational Inheritance Ascriptive after All?

H&L would not endorse the use of a three-variable model to indicate the implications for "ascription" versus "achievement." I both agree and disagree, or rather, I think there is ample room for reasonable differences. They note that education "mediates" such background factors as sibship size, but I regard this as tangential or even distracting (see Rytina 1989). As my initial article in this volume indicates, occupation for me holds center stage. And it is implicit in my approach that the intergenerational correlation is the linchpin around which many partials will turn. Precisely which partials will turn by how much is numerically complex. I have not altered my initial judgment that such baroque considerations would not help clarify my central point.

This central point is, of course, that plausible estimates suggest that the continuity of occupation from father to offspring has been understated by traditional scaling conventions. From this, and other small but consistent changes in the correlations with education, it follows that the direct effect becomes greater. Whether these changes are big enough to worry about is somewhat subjective. This is why my original contrast was with Hout's finding of universalism. Though space does not permit full reexamination here, I have indicated why I see no reason to alter the original estimates and thus still regard "universalism" in the GSS as turning on choice of scale.

H&L do indicate some convergence between our views when they point out that mediation by education is limited in the Blau and Duncan (1967) results, albeit in an effort to attack the novelty of my findings. I have argued elsewhere (Rytina 1989) that the earlier results do not sustain the interpretation of universalism. H&L go so far as to say that "there is no substantial basis for the claim that education mediates more than about half of the effect of father's occupational standing on offspring's" (p. 1702). It is unclear to me why they implicitly dismiss the quantities from the GSS based on the SEI, but with the major amendment of SSIC and the minor amendment of H&L's "more than about half" to the more strict "even as much as half" I agree that there is no warrant for as-

serting that educational mediation is predominant in occupational continuity.

II. SHOULD THE EVIDENCE BE PIGEONHOLED OUT OF SIGHT?

Is empirical occupational mobility peripheral after all? In their attempt to relegate SSIC to a minor footnote of stratification studies, Grusky and Van Rompaey (hereafter G&V) seek to minimize my critique of SEI as an empiricist exercise that lacks analytic warrant. Their analytic procedure is to propose classes for scales and then to fault my efforts when my stated goals do not fit those of the pigeonhole into which they force SSIC. They claim special conceptual status for their favored class of "socioeconomic" but offer little beyond reiteration of an operationalist tautology dignified by tradition. They are, however, quite correct that the original essay emphasized conformity to empiricist canons. This proceeds, as it must, from analytic presuppositions, yet is also deliberately couched in the rhetoric sometimes pejoratively labeled "positivist." A central objective was to display the contrary implications of carrying such empiricism to greater depth than heretofore.

G&V take me to task for failing to mention various related efforts at occupational scaling, although they generously note that this is understandable given my objectives. G&V classify SSIC with Levine's (1990) mobility analysis as an affiliative scale falling within the social network tradition, going back to Laumann (or even, with qualification, Warner). Levine's effort is aggregative in intent and does not use detailed occupation, but it is quite acceptable as a classificatory kin of SSIC.

G&V seem to misinterpret what is involved. They attempt to override my objective of studying inheritance by insisting that, as a consequence of their classification, SSIC is flawed absent isomorphy across such diverse tie types as acquaintance or friendship. That fuzziness obscures the theoretical primacy involved. The social tie that knits occuptions giving rise to a mobility table is sociobiological descent.

Descent is the tie that, along with marriage, generates the web of kinship. (Implicitly, and perhaps regrettably, mobility is most commonly studied as descent from the male line, but here again convention is my guide.) It is descent that marks out the degree to which occupational lines are confined by family, which corresponds to "ascription" in this regard. Occupational continuity within families was taken by such classic writers as Schumpeter and Weber to delimit discrete layers, or strata. A modern-day complication is that long-distance moves are known to be merely improbable, not literally impossible, but with this qualification SSIC directly operationalizes the way that family and occupation inter-

sect to form strata. In this straightforward sense, SSIC can claim to be a direct measure of the intergenerational stratification of occupations.

The SSIC is, then, relational, and it draws on a social bond. But it only obscures matters to group sociobiological inheritance with social networks in general. A better label is "mobility" or "inheritance" scale. Of course, SSIC is hardly the only measure of occupational rank. Occupations, as groupings, implicitly structure how various other attributes are assigned to occupants. Quite unexceptionably, G&V shorthand alternative as prestige, rewards and assets, and socioeconomic standing. Note, however, that the set of distinctions are essentially among the types of data used in construction, that is, face validity.

I found no fault with the theoretical lineage of prestige measurement. Yet it has long been known that such scaling does not exhaust the dependence of objectively scaled offspring rank on origin. Jencks (1990) presents a strong case for the primacy of subjective evaluations of rewards in assessing social mobility rates from the standpoint of popular judgments, but this solves the problem of combining unlike quantities attached to jobs only through an assumption of consensus on evaluation. Briefly put, the objective pattern for occupation is also worthy of sociological attention. One could regard SSIC as an attempt to place objective measurement on a somewhat surer theoretical footing.

The alternative is socioeconomic standing. G&V perceive a theoretical tradition in various pragmatically guided attempts sharing the concern that something more than pure economic status was a useful way of marking off occupational ranks. I see here only an operational convention that fudges two central issues. Education is, at best, a proxy that does not exhaust the broad conceptual warrant of *socio*. The rules for combining the several dimensions are inevitably arbitrary.

G&V claim to find in sources such as Hodge (1981) a middle-range theory or concept underlying socioeconomic scaling. I found Hodge's commentary far more skeptical. He stated that "an even greater difficulty with the socioeconomic interpretation of Duncan's SEI scales stems from the analytic status of the concept of 'socieconomic level.' As far as we can see, it has none" (1981, p. 406). Hodge noted that Duncan acknowledged that alternative "reasonable procedures . . . are easily proposed" (Duncan 1961, p. 119). He underlined the classic operationalist tautology: "Socioeconomic status is what socioeconomic scales measure; there is no underlying analytical concept to which we can refer a proposed indicator of socioeconomic status to decide whether it is well or ill-conceived or to assess how it might be improved" (Hodge 1981, p. 407).

The SSIC can be seen as a solution to Hodge's dilemma. While no absolute critique of SEI is feasible, this scale is entrenched in the linear

modeling of intergenerational mobility. Such models invoke the implicit assumption that occupations that are equal on SEI lead to equal outcomes. This assumption is empirically false. The SSIC is an effective corrective. Occupations are equated by SSIC precisely insofar as they are equivalent in average intergenerational mobility experiences. Levine (1990) called this "behavioral equivalence." His context was slightly different, but mutatis mutandis, his behavioral equivalence is an attractive name for an attractive property enjoyed by SSIC.

Behavioral equivalence is, unquestionably, an empirical property or standard, but one that corresponds with the widely held analytic notion that intergenerational layering is a fundamental feature of stratification. It provides the backdrop or baseline for derivative efforts to understand what variables and processes account for or condition the empirical pattern of occupational continuity across generations.

While G&V generously allow that my case for interpreting such results as reflecting "assets and advantages embedded in occupations" makes for "compelling reading" (p. 1721), they question whether all mobility exchanges are reflections of a pure socioeconomic hierarchy. They compare this to the assumption of perfect markets of an ideal-typical economist. In a way, the comparison is apt. My aim was to define an analytically simple starting point that lent itself to a straightforward ideal-typical verbalization. Unlike that of the economist, my approach implicitly incorporates such market imperfections as discrimination on social origins.

G&V are not persuaded that such results are "really" socioeconomic. The only possibilities they raise would trouble their "ideal-typical" economist more than a sociologist. G&V would single out effects of barriers due to industrial, regional, and ethnic labor markets as "idiosyncratic." Since SSIC turns on averages for whole occupations, at issue are wholesale departures from SEI linearity such that persons are confined to restricted ranges of occupations for generation after generation.

On their account, such effects of ethnic or industrial labor markets are something distinct from "socioeconomic" stratification. This is true but uninteresting if socioeconomic status is by definition what socioeconomic scales measure. It is true but ideologically laden if socioeconomic stratification refers to an ideal pattern purged of distorting market imperfections. A more attractive alternative is to eschew the narrow face validity construal of G&V and admit such systematic departures under the conceptual umbrella of "socioeconomic." With respect to SSIC, this is but a guiding interpretation for future research into exact mechanism(s), but there is no barrier in principle to regarding systematic discrepancies in mobility as reflecting socioeconomic causes and effects.

It remains true that SSIC is conceptually and empirically distinct from

other scales. Summarizing mobility and its counterpart of ascription is not the same as averaging income and education. But the pattern that results has real, that is, empirical, implications, which are manifest even if SEI or prestige is taken as the criterion.

G&V are in error when they suggest that SSIC rests on an arbitrary criterion variable (p. 1724). As the demonstration using rough and ready scaling showed, the criterion is quite nearly equivalent to mobility chances as measured by SEI. This criterion derives from the presumed adequacy of the linear status attainment representation. This is hardly arbitrary, whatever its ultimate status as analytic presupposition.

G&V speculate about alternative indices for ascription other than occupational ascent and descent. They note that coding occupation by mean educational attainments of descendants would beef up educational mediation. Much the same might occur through coding as measured intelligence or physical stature of average offspring. But such coding "measures" offspring occupational advantage as whatever is best for the next generation in various regards. What is mediated is not adult occupational advantage, but something more like educational advantage for the grandchildren. The SSIC, by contrast, records the occupations obtained by the children from the most advantaged families as best.

Once again, there are other ways to scale occupation, implicitly as many as attributes that could be averaged. Many of these may be educationally mediated; prestige is an analytically interesting example. But only SSIC draws on occupation alone without restriction to particular attributes. It describes, subject only to the qualification of linearity, how the division of paid labor is intergenerationally reproduced. Such a statistical summary is unquestionably an idealization, but the type to which it corresponds is the separation of occupations into strata by ascription. And it shows that ascriptive occupational rank, "as it is to be seen" by disaggregating to detailed occupation, is but weakly mediated by education.

REFERENCES

Blau, Peter M., and Otis Dudley Duncan. 1967. *The American Occupational Structure* New York: Academic.
Data Program and Library Service. 1983. *Occupational Changes in a Generation, 1973: A User's Guide to the Machine Readable Data File.* Madison, Wis.: Data Program and Library Service.
Davis, James A., and Tom Smith 1986. *General Social Survey 1972–86: Cumulative Data* (MRDF).NORC ed Chicago: NORC. Distributed by Roper Public Opinion Research Center, New Haven, Conn.
Duncan, Otis Dudley. 1961. "A Socioeconomic Index for All Occupations." Pp. 109–38. *Occupations and Social Status,* edited by A Reiss. New York: Free Press

Duncan, Otis Dudley. 1979. "How Destination Depends on Origin in the Occupational Mobility Table." *American Journal of Sociology* 84:793–803.

Featherman, David, and Robert Hauser. 1976. "Prestige or Socioeconomic Scales in the Study of Occupational Achievement." *Sociological Methods and Research* 4 (May)· 403–23.

———. 1978. *Opportunity and Change.* New York: Academic.

Featherman, David L., F. Lancaster Jones, and Robert M. Hauser 1975. "Assumptions of Social Mobility Research in the U.S.: The Case of Occupational Status." *Social Science Research* 4:329–60

Ghiselli, Edwin E., John P. Campbell, and Sheldon Zedeck. 1981. *Measurement Theory for the Behavioral Sciences* San Francisco. Freeman.

Goodman, Leo A. 1979a. "Multiplicative Models for the Analysis of Occupational Mobility Tables and Other Kinds of Cross-classification Tables." *American Journal of Sociology* 84.804–19.

———. 1979b. "Simple Models for the Analysis of Association in Cross-classifications Having Ordered Categories." *Journal of the American Statistical Association* 74:537–52.

———. 1981a. "Association Models and Canonical Correlation in the Analysis of Cross-classifications Having Ordered Categories." *Journal of the American Statistical Association* 76:320–34.

———. 1981b. "Association Models and the Bivariate Normal for Contingency Tables with Ordered Categories." *Biometrika* 68.347–55.

Hodge, Robert M. 1981. "The Measurement of Occupational Status." *Social Science Research* 10:396–415.

Jencks, Christopher. 1990. "What Is the True Rate of Social Mobility?" Pp. 103–30 in *Social Mobility and Social Structure,* edited by Ronald L. Breiger. New York: Cambridge University Press.

Jencks, Christopher, Susan Bartlett, Mary Corcoran, James Crouse, David Eaglesfield, Gregory Jackson, Kent McClelland, Peter Mueser, Michael Olneck, Joseph Schwartz, Sherry Ward, and Jill Williams. 1979. *Who Gets Ahead? The Determinants of Economic Success in America* New York: Basic.

Lawley, D. N. 1959. "Tests of Significance in Canonical Analysis." *Biometrika* 46:59–66.

Levine, Joel H. 1990. "Measuring Occupational Stratification Using Log-linear Distance Models." Pp. 208–24 in *Social Mobility and Social Structure,* ed. Ronald L. Breiger. New York: Cambridge University Press

Rytina, Steve. 1989. "Life Chances and the Continuity of Rank." *American Sociological Review* 54:810–19.

———. 1991. "Return to Normality: Simple Structure for the Mobility Table and the Symmetric Scaling of Intergenerational Continuity." Manuscript. McGill University, Department of Sociology, February.

Stevens, Gillian, and David Featherman. 1981. "A Revised Socioeconomic Index of Occupational Status." *Social Science Research* 10:364–95.

Book Reviews

Landscapes of Power: From Detroit to Disney World. By Sharon Zukin. Berkeley and Los Angeles: University of California Press, 1991. Pp. xii + 348. $24.95.

Richard Sennett
New York University

Sharon Zukin's *Landscapes of Power* makes bold use of visual evidence to advance a sophisticated view of modern society. It is a beautifully written, intelligent, and yet deeply frustrating book. In brief, Zukin is concerned with how modern market forces of capital, labor, and consumption intersect with experiences of place. The modern market economy Zukin takes to be transnational in its scope, constantly shifting in its locations; its consumerism is less about the acquisition and retention of solid goods than the pursuit of dreams. The modern market is constantly at work to erode the integrity of territorial places; the transnational mobility of labor and capital make cities into sites for economic and social activity that are beyond the power of any particular city to control. These same forces disorient and fragment place in another sense, that of the images people use to place themselves in relation to each other; this social disorientation occurs through the ever-shifting fantasies of consumption. Yet, she argues, people's need for turf and for a coherent image of their placement in society are compelling social needs. The interacting, conflicting relations between market and place produce what she calls "landscapes of power." These are visual forms that simultaneously represent how the economic and the social intersect and mediate their dissonances.

Landscapes of Power elaborates this theme through five case studies. The first two are of traditional industrial-urban forms: the company town of Weirton, West Virginia and the industrial city of Detroit. The emphasis in both is on how the modern market dislocates traditional urban forms; in these case studies the focus is on place as territory. A transitional case study is of Westchester County, New York, a suburb that has become largely a corporate headquarters; Zukin explores the tension between the traditional, pastoral imagery of the suburb as an escape from the city and the territorial needs and architectural forms dictated by big corporate headquarters. The last two case studies focus on the interaction of consumer economics and the sense of social relationship; one depicts the landscape of gentrification, particularly in New York, the other, the

fantasy of pure consumption pleasure, is embodied in the Disney World parks and their relations to Miami and Los Angeles.

Landscapes of Power is a work of what might be called visual sociology. Rather than using verbal evidence like questionnaires, it asks "What does society look like?" The question is meant literally. In a way, visual sociology has a long history; the early ethnographies of the Chicago School were filled with the evidence of the eye, and Zukin's great gifts for describing what she sees are reminiscent of those ethnographies, such as Harvey Zorbaugh's *The Gold Cost and the Slum* (Chicago: University of Chicago Press, 1976). In another way, her work, like that of others now writing in this vein, marks a sharp break from the past.

The Chicago ethnographers took visual evidence as simple, straightforward representations of "deeper" social forces like class or ethnicity. The visual sociology practiced by Zukin, like that by David Harvey, Pierre Bourdieu, and myself, takes visual evidence to be more problematic. This is partly due to what we have learned from art historians about the omissions, disguises, and feints involved in the act of visual representation. Some of the problematic nature of visual evidence, however, comes from its very sociological potency and character. The visual constructs the social: images define "us" and "them"; gender relations are constructed from bodily appearances; the legitimacy of power, as the Romans well knew, turns on the capacity to make a compelling architecture. The visual constructs the social, yet the process by which this happens is, at the level of analysis and theory, obscure.

Sharon Zukin's book runs aground on this problem. She argues, as she puts it at the end of her study, that "space is a major structuring medium," (p. 268) creating the market economy as well as being created by it. She wants to do something in visual terms that is akin to Anthony Giddens's work on structuration. Yet her five case studies do not accomplish this; the case studies treat the visual as a consequence of the economic, and one finishes the book with a strong sense of a gap between description and theory. Indeed, her book fails to justify the use of the word "landscape" in its title; classical landscapes, as the art historian Svetlana Alpers has shown in her book *The Art of Describing* (Chicago: University of Chicago Press, 1983), are "corrective fictions," images of the world as it could be and should be. *Landscapes of Power* conveys little about how power might look differently; there is little theory about the powers of vision and therefore those of re-vision.

I do not mean to suggest that Sharon Zukin has failed to write a good book. On the contrary, it is the very strength of her capacity to observe that poses so forcefully the problem faced by all those engaged in this kind of work: how to understand the social consequences of the human power to see. This book is a distinguished work in progress in a larger, unfinished project.

The Saturated Self: Dilemmas of Identity in Contemporary Life. By Kenneth J. Gergen. New York: Basic Books, 1991. Pp. xiv + 295. $24.95.

Philip Wexler
University of Rochester

For more than 20 years Kenneth Gergen has been both an important psychologist of the self and a historical and cultural critic of mainstream American psychology. The effort to understand the self has increasingly fused with Gergen's relativizing of psychology to create a linguistic, if not literary, orientation to vocabularies of self, consciousness, and being. *The Saturated Self* fulfills this tendency by contextualizing both the self and its cultural relativism in a larger historical social frame that Gergen identifies as postmodernism.

The argument of the book is that the languages, being, and consciousness of self are cultural expressions and that, while there are residues of the prior cultures of romanticism and modernism in selfhood, a new relational, "populated" self emerges with the culture of postmodernism. For Gergen, the new culture is expressed both within academic crises about the certainties and even possibilities of knowledge, and in the daily lives of nonacademic selves. The proximate source of the critical shift is, according to Gergen, to be found in communicatory technologies which, by social saturation, fundamentally alter the terms of knowledge and identity.

The postmodern self described appears at first to be simply a dissolution of the self. But, the disappearance of earlier typical selves leaves not sheer emptiness, but a relational self. The new social world of media- and communicatory-technology saturation and intensification of relatedness (albeit distant, but putatively, a no less powerful intimacy) induces a self with a potential for unlimited multiplicity—"multiphrenia"—on the one hand, and, on the other, opens the way to a new relational self that results from the "growth in social connectedness" and offers the promise of a new, postmodern community.

Around this central theme, Gergen interweaves (in an appropriately postmodern style) examples and anecdotes from his own and friends' experiences and summaries of and references to a wide range of works in social and literary theory, philosophy, and cultural commentary. For the novice in the wild and woolly world of an enormous literature in poststructuralism, pragmatism, hermeneutics, anthropology, and cultural criticism that Gergen partially appropriates as a critique of modernism and romanticism and stitches together under the sometimes forced rubric of postmodernism, this can be a good, light appetizer to a more serious and systematic meal of the same fare. What is particularly valuable, however, is the effort to review selections from this literature from the vantage point of an engaging assertion about a historical transformation of the self.

To many American psychologists, Gergen's book, like his earlier work, will seem heretical; a minority will follow Gergen's lead onto a bigger screen, one that is increasingly legitimatized for them under the heading of "context." Sociologists who have been unaware of recent social theoretical work outside of conventional empiricist paradigms may find this book a useful introduction. Some social theorists may think, to use Geertz's phrase, the "blurring of genres" here is still a little too postmodernist for their analytical taste, or, more tellingly, has little social news for them.

One of my reservations is that, despite a brief disclaimer, the social fabric from which this antistory story is cut is a bit too narrow. Gergen's examples are the stuff of the lifeworld of a segment of the upper-professional middle classes; the content of the communicatory and relational revolution he describes border on cultural arrogance—within a wider "context" of a social structure of inequality and immiserization. Beyond this social structural reservation, I have an empirical one, not unrelated to the existence not simply of "multiple," but stratified lifeworlds. In *our* studies of everyday life among contemporary youths of different social strata (*Becoming Somebody*, Falmer Press, 1992), *we* do find a dissolution of earlier social and self certainties. But the dissolution is not welcomed easily. On the contrary, across class differences in cultural substance, everyday life is characterized by an evermore difficult struggle to establish a secure and integral self-identity in the face of social contradiction and disintegration. The "dilemmas of identity in contemporary life" that still need to be described are of many real people in real social institutions. I hope that Gergen's book will encourage such social analyses.

Commonplaces: Community, Ideology and Identity in American Culture. By David M. Hummon. Albany: State University of New York Press, 1990. Pp. xv + 236. $49.50 (cloth); $16.95 (paper).

Wendy Griswold
University of Chicago

Ideology is the process whereby what is historical—produced in time and place, artificial, and in service of particular interests—comes to appear natural. Theorists of ideology have contended that the basic categories through which humans organize their social worlds, categories like male/female, human/animal, normal/deviant, are ideological constructs, though nonetheless powerful for all their arbitrariness. Nowhere has this argument been made to better effect than in discussions of community. Forsaking the 19th-century assumption that community involved a natural intimacy built on the bedrock of kinship, social science has moved to a social constructivist approach. As Benedict Anderson puts it, "All

communities beyond face-to-face are imagined"; they differ not in their truth or falsity, but in "the style in which they are imagined."

In *Commonplaces: Community, Ideology and Identity in American Culture,* David M. Hummon is tilling this vineyard, attempting to discern the style in which contemporary Americans imagine their communities. He is concerned with delineating the exact nature of the ideology as it currently operates or, in other words, with assessing if our "commonplaces" about communities of residence, their characteristics, and their desirability, are in fact commonly held. According to Hummon, Americans' experience of place is available to consciousness only as mediated by community perspectives, that is, the understandings of place—city, suburb, small town—given by our shared culture (or, "commonplace" knowledge about common places). No one comes to a residential setting cold. Children pick up the story of the city mouse and country mouse and its implications about urban versus rural life, for example, even before they are clear about in which type of setting they happen to live. Hummon's metaphor of perspective indicates that cultural commonplaces shape how we see both the place where we stand and those distant places on the horizon (hick towns, the burbs, the Big Apple) that attract or repel us.

Hummon interviewed residents of a city (San Francisco), two suburbs (one affluent, one not), and a small town in northern California to find out what the various community perspectives looked like. His empirical aim was to interpret these interviews to show "how a shared, cultural perspective, drawn from a community ideology, transforms community belief into community commitment" (p. 38). He found that discourse about place is systemic—one cannot talk about living in a small town without talking about "the city," and vice versa—and beliefs about types of community assume competition among these types. Indeed, a key theoretical point is that community perspectives do not just reflect experience in a place, but constitute an argument about *places.* Such arguments may be folded back into identity, both of oneself ("I'm a city person") and of others ("She's a small-town girl at heart").

One intriguing point of ideological slippage is the blurred cultural identity of the suburbs. While Americans have well-defined perspectives regarding cities and small towns, suburbs tend to be invisible from the outside. City folk see them as lumped with small towns and everything else bush league, while small town folk see them as just part of the crime-ridden urban sprawl. More surprising, suburbanites themselves are a little vague about the kind of place in which they find themselves. Some suburbanites view themselves as living in small towns, others claim to be living in a city. Still others, but only a minority, are more specific and assertive, maintaining that the suburbs have the best of both worlds (outsiders tend to assume they have the worst).

Commonplaces is filled with gems, some of which are no less bright for being commonplace. Small-town folks are more satisfied than others, do not recognize that suburbs exist, fear cities, and congratulate themselves

1753

all over the place on their neighborliness. City folks disdain suburbs and small towns as stifling, boost their own cities but not always city life in general, equate all that neighborliness with nosiness, are often vaguely dissatisfied, and think urban crime problems are much overrated even as they pat themselves on the back for the nimbleness with which they avoid muggers. They also value cities for being liberal and tolerant, zones of freedom, although on this point Hummon should have paid more attention to the fact that his particular city was San Francisco; one suspects that urbanites from Kansas City or Detroit would not share this perspective. As for suburbanites, a minority ("suburbanists") promote the best-of-both-worlds line, emphasizing the suburbs as ideal locations for family life, while the majority, regarding themselves as actually living in either a small town or a city, sound like city folk or small-town folk when they talk of the advantages of their particular place of residence. In other words, small towners are complacent, city dwellers frenetic, suburbanites fuzzy; some of this is just what we always suspected.

Persuasive as it is about the systematic ideology of place that contemporary Americans share, the book does not fully achieve Hummon's larger objectives of explaining the link between American culture and community ideology (he simply asserts it) or of connecting community beliefs with community commitment. There are no measures of commitment at all, in fact, although clearly it's one thing to say "I'm a neighborly small-town person" and another thing to vote for a local school bond issue if you're childless. Community perspectives do allow the formation of community accounts that "make sentiments sensible," as when a devout urbanist remarks that "People in small towns are more friendly because they've got nothing else to do [chuckles]" (p. 127). But the relation of such sentiments to behavior is never explored. By failing to distinguish commitment from sentiment, *Commonplaces* implies that what people claim to feel and how they act are linked directly and obviously. Hummon, like all sociologists, knows better.

Mea Culpa: A Sociology of Apology and Reconciliation. By Nicholas Tavuchis. Stanford, Calif.: Stanford University Press, 1991. Pp. ix + 165. $29.50.

Suzanne Retzinger
University of California, Santa Barbara

Mea Culpa is an important book. Tavuchis considers apologies between individuals, individuals and groups, and between groups. He addresses the basic sociological questions of membership, deviance, and conformity in a new way. Tavuchis describes the intricate process of apology as a dyadic interaction between offender and offended, where remorse and forgiveness are central. A common thread runs between the macro- and

microlevels of analysis; apology is not only about reparation of transgressions; it reveals the core of human relationships.

Tavuchis's analysis is broad and interdisciplinary, drawing from sociology, philosophy, sociolinguistics, social psychology, anthropology, philology, law, and religion. It is micro and macro, interactional and structural. The microprocesses of apology reveal the structure of groups; macroanalysis exposes individual process. He utilizes verbatim texts (from newspapers, novels, letters, press conferences) to develop his theory.

Tavuchis is particularly brilliant on the work of Goffman. Unlike Goffman, Tavuchis goes beyond the individual to the relational. Apology, he says, is not a single action but a delicate sequence of events between persons; he shows how apology is remedial, by leading to forgiveness and reconciliation. Tavuchis goes further than Goffman's behavioral analysis, he deals with feelings: one must not only say one is sorry, one must *feel* sorry.

Apology is set apart from excuses, defenses, justifications, and accounts, which cover and hide the self, shift responsibility away from self, distance persons from each other, and undermine the bond. Apology involves risks and self exposure, one takes responsibility for one's actions; apology leaves open the possibility of repair and solidarity, a prelude to reunion.

Wrongdoing cannot go unnoticed without compromising the relationship, yet cannot be undone. Apology is a paradox. A crucial observation which Tavuchis makes is that although apology cannot undo what has been done, it has the "power to rehabilitate the individual and restore social harmony" (p. 9). In this way it defies exchange and rational choice theories. Apology speaks to something larger than the particular offense and cannot be understood in terms of expediency.

Ultimately, apology is about the reparation of social bonds. Apology sets things right, not by transforming the material world, but by its power to transform emotions and restore bonds. Tavuchis shows how apology is ultimately and intricately tied to community and to social order.

Apologetic acts are interactive, the offender must elicit forgiveness from the offended. For an apology to occur there must also be a valued bond between offender and offended. To elicit an apology the offended communicates to the offender the injury. Offender must feel sorry, state his or her offense, take full responsibility for the wrongdoing, and verbally express regret, with the promise that the offense will not reoccur. Timing and sincerity are of the essence. Forgiveness is the only way to alleviate the shame of the offender and save the relationship. The offender puts his or her self at the mercy of the offended.

Some of the mystery of apology might be removed if the act is described nonverbally as well as verbally. Although Tavuchis stresses the importance of the verbal act, the manner (tone, pitch, cadence, timing, rhythm, eye gaze, etc.) in which an apology is elicited and given has at least equal importance with the actual spoken words; manner conveys emotion.

Tavuchis mentions certain emotional states in the interaction between persons, but does not place enough importance on them. In apology where self-exposure and admission of defenselessness are crucial, shame particularly plays a role. Apologizing and forgiving may involve a delicate dance of shame and the ability to acknowledge this painful emotion by both parties. The offender must be able to feel and acknowledge his own shame to self, transform this shame into guilt and remorse, and admit personal failings. The offended must be able to communicate injury in a respectful manner, rather than retaliate. The dance of emotion might help explain why some parties prefer to escalate anger and hatred, rather than apologize and forgive (*Violent Emotions* by S. Retzinger [Newbury Park, Calif.: Sage, 1991] and *Violence and Emotion* by T. Scheff and S. Retzinger [New York: Lexington, 1991]).

After reading *Mea Culpa*, I am struck with the question of why we see so little real apology; many more excuses are offered than apologies. In a society where the individual is deified, where bonds and social emotions are denied, apology is difficult; persons have to have some sense of community or solidarity for apology to occur. Where alienation reigns, apology is problematic. Although apology can remove threats to the bond, risking exposure can also feel like a threat to an already weakened bond. Ability to apologize reveals the level of alienation or solidarity in any given group.

Tavuchis's work has many parallels with that of John Braithwaite (*Crime, Shame and Reintegration* [New York: Cambridge University Press, 1989]). Braithwaite's formulation of reintegrative shame also contains a process of repair. With the admission of the wrongful deed and proper apology, the offender is forgiven and reintegrated into the community. When forgiveness is absent the offender is segregated and stigmatized. Tavuchis's work provides examples of criminal apology that may prove important in the judicial system.

Mea Culpa is a valuable contribution to social science. Tavuchis has made a powerful start, but I am left wanting to know more: What types of bonds enable apology and forgiveness? What is the role of shame and guilt? Do altruistic societies (like Japan) produce members with greater ability to apologize than those of anomic societies? What is the role of solidarity in an apology between the many and the one? Like any wonderful book, *Mea Culpa* leaves the reader with both new insight and new questions.

The Egalitarians—Human and Chimpanzee: An Anthropological View of Social Organization. By Margaret Power. Cambridge: Cambridge University Press, 1991. Pp. xviii+290. $44.95.

Theodore D. Kemper
St. John's University

If social science has contributed anything to human knowledge, it is that social structure and culture can override, sometimes even determine, what appear to be biologically innate properties of human beings. Margaret Power's comparative study lends interesting new authority to this truth.

Essentially, Power argues that because human hunter-gatherers and chimpanzees in the wild share a similar ecological niche, their social organization is remarkably similar. The qualifier, *in the wild,* is significant, inasmuch as the dominant paradigm in chimpanzee studies today derives from the later work of Jane Goodall, who reports that the animals are strongly territorial, aggressive, and dominance-seeking. Whereas Goodall's analysis might support a theory of phylogenetic continuity for similar, *biologically* inherent, agonistic qualities in humans, Power's important contribution is to show that Goodall's conclusions may rest principally on the "unnatural" environment that Goodall herself created for the apes in order to facilitate observation of their behavior.

When Goodall began her naturalistic studies of chimpanzees in 1960 in the Gombe National Park area of Tanzania, she was a distinctly *nonparticipant* observer. After some years of patiently tracking apes over large areas, Goodall discovered that she could lure the animals into a more or less permanent presence around her camp, thereby improving opportunities to observe social interaction, by baiting the camp with supplies of bananas. Indeed, this was an inspired notion. According to Power, it worked too well.

Power maintains that the change that Goodall engineered in the food supply warped the chimpanzees' conduct and social organization more or less permanently. Power pursues the argument by examining the differences between Goodall's observations prior to the artificial feeding regimen and the subsequent findings. Goodall herself does not rely much on the results of her early work.

Power argues that, like human hunter-gatherers, chimpanzees in the wild roam widely, rarely confronting each other in direct competition over food. Goodall's artificial feeding, practiced from 1964 to 1968, introduced direct competition among the apes for the first time. Bunched around the feeding boxes and often frustrated by not obtaining the bananas (which were doled out according to specified schedules), the animals began to engage in more intense forms of competitive, aggressive, and threatening behavior than was known to occur in the wild. Domi-

nance now became functional, since it insured a larger supply of bananas once the feeding boxes were opened.

Power contends that apes, like humans, acquire a substantial amount of useful knowledge through socialization. After the emergence of the unwonted patterns of aggression and dominance-seeking, which invaded chimpanzee society (even away from the feeding station), young male chimpanzees learned to model the techniques that manifestly brought more food to the elder males (more access to females, too, since dominance also gained that advantage). Females learned to wean their offspring earlier than before so that they could better protect themselves from aggressive attack. This new practice led to early frustration of dependency needs, producing yet another instigation to aggression.

Although Goodall (*The Chimpanzees of Gombe, Patterns of Behavior* [Cambridge, Mass.: Harvard University Press, 1986]) has acknowledged that the artificial feeding led to a dramatic rise in aggression, she maintains that when it was largely, but not completely, abandoned in 1969, aggression declined. Yet, says Power, it was too late. A crucial change in chimpanzee social organization had been initiated by the feeding process, and a return to the earlier, natural pattern was precluded. There is one very chilling indication that it did not return, and this is among Power's most telling points.

In the course of the normal fissioning that takes place in chimpanzee society, a small group separated from the main group and established another home range for itself in neighboring territory. Over the course of about 10 years, all members of this community disappeared, with every indication that they had been killed by animals from the original group. From this event, Goodall derived that chimpanzees possess the protohuman qualities that underlie the capacity to wage organized warfare. Power, on the other hand, maintains that chimpanzees in the wild do not engage in such violent conduct, ordinarily retreating when they encounter strangers from neighboring groups. Power's point is that the "natural" habitat of chimpanzees does not dispose them to genocide. Nor, by extension, would the natural habitat of humans.

Margaret Power takes Rousseau seriously. For her, there *is* a state of nature, it is similar for both humans and chimpanzees, and in it conditions are more, rather than less, egalitarian and irenic. Neither aggressive behavior nor dominance over others is especially functional. Since situational demands for leadership change from day to day, there is no permanent hierarchy. Males and females have equipotential for initiating and declining sexual invitations, and gender roles are differentiated mainly with respect to the degree of intimacy in the nurturance of offspring. Power implies that the parallels between chimpanzee and early human forager groups allow us to treat at least some phylogenetically earlier data as indicative of the "natural" human condition. Indeed, at points, Power so interweaves the data on animals and humans that only close attention reveals which species is the subject. Ultimately, we are left

with the sense that the animals we have seen in the several excellent television documentaries about Goodall's work may be as distant in crucial respects from their own species mates in the wild as we are from our species mates in the few foraging communities still left on earth.

Social Evolutionism: A Critical History. By Stephen K. Sanderson. Cambridge, Mass.: Basil Blackwell, 1990. Pp. xviii + 251. $39.95.

Mark A. Shields
Georgia Institute of Technology

Sociology emerged as a philosophically and historically oriented discipline that sought to provide an evolutionary account of the nature of industrial society and modern culture. Virtually all of the discipline's founding fathers—Comte, Saint-Simon, Spencer, Marx, Durkheim, and Weber— were to varying degrees evolutionists. Yet sociology has always had a deep intellectual ambivalence, and periodic animus, toward thinking in evolutionary terms. Within the past two decades or so the animus against evolutionary theory has intensified. A majority of sociologists now would probably reject the very *idea* of social or cultural evolution as empirically implausible, and regard evolutionary theory as an impertinent expression of discredited teleological thinking. Among other damaging objections to evolutionary theory, critics have argued that its explanatory scope is too sweeping and overgeneralized, lacks empirically testable propositions, is afflicted by outmoded presuppositions from the philosophy of history, is ensnared by value-laden notions of progress, and is tainted by facile analogies and metaphors drawn from evolutionary biology.

Stephen Sanderson's wide-ranging exposition and critique of evolutionary theories from the mid-19th century on weighs these and other objections, but argues persuasively that the antievolutionist critique has often been rash and precipitous—certainly overgeneralized. He sets out "to expose the considerable mythology" about social evolutionism and "to show that certain versions of evolutionary theory are intellectually defensible" (p. 223). Doing so, Sanderson presents concise, informative summaries of an impressive range of classical and contemporary evolutionary theorists, including Spencer, Morgan, Tylor, Marx and Engels, Childe, White, Steward, Sahlins, Service, Carneiro, Lenski, Harris, and Parsons. An excellent chapter on the relationship of evolutionary biology to social evolutionism examines the concepts of adapation, differentiation, complexity, and progress. A chapter on contemporary antievolutionism—focusing on Nisbet, Giddens, and Mandelbaum—adumbrates the major persistent challenges posed to evolutionists by thoughtful critics. As Sanderson makes clear, his book is not "an account of the social and cultural conditions that have [influenced] the reception of evolutionary theories" (p. 7), nor does it propose a new evolutionary theory. His final

chapter, nonetheless, is a schematic, bare-bones outline of what needs to be done (and undone) in reconstructing social evolutionary theory.

All in all, *Social Evolutionism: A Critical History* is the best available single volume examination of this complex and contentious theoretical tradition. Its numerous strengths are impressive. It is concise but comprehensive in theoretical coverage, intelligent in addressing a number of crucial distinctions, such as that between evolutionary and evolutionist theories, carefully documented with primary and secondary sources, generally fairminded in exposition and criticism, and capably written with clarity and coherence that is rare in scholarly treatments of evolutionary theory. This book is a worthy addition to the *Studies in Social Discontinuity* series edited by Charles Tilly. Scholars interested in social evolutionary theory will want it on their shelves as a standard reference work.

Its shortcomings are also apparent. Sanderson dogmatically embraces sociological materialism, and his writing at times is tiresome because of his preoccupation with slotting complex ideas and theorists into an idealist-materialist dichotomy. Sanderson thoughtfully singles out the antiprogressivist cultural materialism of Marvin Harris as the most promising social evolutionary theory, while his severest criticism is directed at Parsons's alleged developmentalist teleology and causal idealism. While he does an excellent job of discussing Parsons's evolutionary theory and its key shortcomings, he needlessly caricatures Parsons's logic with statements like: "Parsons is really making the teleological argument that evolution occurs *in order to bring about advancement*" (his emphasis; p. 116). Sanderson's generally balanced coverage of important theorists is missing in the scant attention he gives to the "neoevolutionists" Luhmann and Habermas (pp. 127–28). He dispenses with Habermas in just over one page as essentially a neo-Parsonian developmental idealist. Habermas, in particular, deserves more sophisticated attention than this in a volume on social evolutionism. Arguably, no contemporary evolutionary theorist presents more complex paradoxes and challenges to prevailing intellectual dichotomies. Another problem is that Sanderson apparently thinks that social evolutionary theory can aspire simply to construct a kind of value-neutral theoretical edifice of "ordinary causal explanations" (chap. 2), somehow transcending the controversial normative and theoretical presuppositions that ignite most of the perennial big debates in social theory.

Sanderson is convinced that a comprehensive theory of sociocultural evolution is within our reach, if not our grasp: "world history can still be—indeed, should be—understood in evolutionary terms" (p. 223). In the final chapter, he proposes nine desiderata for a "formalized, propositional theory of sociocultural evolution" (p.223), and suggests that they would need to apply to "three fundamental transformations in human history" (p. 224): the Neolithic Revolution, the rise of civilization and the state, and the transition from feudalism to capitalism. Seven of the nine desiderata are, I believe, outstanding criteria for evaluating existing

theories and for building better ones. The other two (desiderata 3 and 5) are dubious. It is not at all clear that an adequate theory of social evolution should "make the individual . . . the unit of adaptation," as Sanderson argues, yet it is clear that "some abstract social system" (p. 224) will not work either. Rather, in principle, multiple levels of adaptation (and maladaptation) in differentiated sociocultural systems need to be specified and interrelated. Another desideratum—"adopt a multidimensional materialist conception of explanation emphasizing the causal priority of demographic, ecological, technological, and economic factors (. . . allowing for a certain amount of 'superstructural feedback')" (p. 223)—is also dubious (and certainly not "multidimensional" in the fullest sense), albeit consistent with his presuppositional commitments to sociological materialism.

Can social evolutionism as a theoretical project be defended today? Sanderson obviously believes it can, and he does a commendable job of explaining why. Despite the book's strengths, however, I doubt that it will dissuade harsh critics or embolden ardent proponents of social evolutionism. Yet I am sure critics and proponents alike will find much to appreciate in this fine work.

Democracy and the Market: Political and Economic Reforms in Eastern Europe and Latin America. By Adam Przeworski. Cambridge: Cambridge University Press, 1991. Pp. ix + 210. $39.50 (cloth); $12.95 (paper).

Michael Hechter
University of Arizona

Recent political events in Eastern Europe and the Soviet Union have been so dramatic and so unexpected that many have been led to wonder anew about the limits of modern social analysis: that phenomena of such magnitude could occur with so little forewarning is discouraging. Now that Communist regimes have crumbled everywhere on the European subcontinent, what will happen next? Will Eastern Europe be able to institute enduring democratic institutions and embark on a period of sustained economic growth, as the optimists hope? Or will capitalism in the East come to resemble the politically unstable and economically stagnant kind of social system found in much of Latin America, as pessimists fear? These are the questions posed by *Democracy and the Market*.

Few scholars are as well equipped to offer an answer as political scientist Adam Przeworski. An exiled Pole who commutes between Paris and Chicago, Przeworski has broad substantive knowledge of Latin American and Eastern European societies, as well as a sophisticated understanding of general theory. How rare a combination of inductive and deductive skills this is among contemporary scholars. So, despite the evident haste of its preparation (this is less a book than a collection of loosely related articles), we expect a great deal from this volume.

Przeworski begins by asking the central question about the durability of democracy: "How does it happen that political forces that lose in contestation comply with the outcomes and continue to participate rather than subvert democratic institutions?" (p. 15). This is a particular instantiation of the far more general problem of social order. The requisite compliance can be explained in terms of one of three different kinds of mechanisms: (1) by spontaneous self-enforcing equilibria derived from the theory of repeated games, (2) by the establishment of some kind of social contract, or (3) by processes leading to the internalization of norms. Przeworski avers that social contracts beg the question because their enforcement requires the kind of compliance that he is setting out to explain and that reliance on internalized norms is unnecessary. Repeated game theory is all that is left on his list; it shows that cooperation among rational egoists can be spontaneously enforced in systems with decentralized authority. As Przeworski himself admits, however, this claim is extremely contingent (for a current assessment of the utility of repeated game theory for questions of this sort, see the special issue on game theory in the social sciences, *Rationality and Society* 4, 1 [1992]).

Przeworski's use of game theory is highly idiosyncratic. His actors are not the game theorists' usual atomistic individuals—(here dismissed as "ideologically derived and patently unreasonable . . . 'individuals' who in addition appear homogeneous in that they have available to them the same strategies and often the same payoffs" [p. 38, n. 48]). Some will be comforted to learn that his players are existing and potential *collective actors,* such as the military or the bourgeoisie, who are embedded in a given institutional framework (p. 26, n. 29).

According to Przeworski, democracy will be self-enforcing when all the relevant political forces have some specific minimum (but unspecified) probability of doing well under the new institutional order. But this probability varies among different collective actors. For example, unions are more likely to be favorable than the military on this account. Przeworski echoes themes from Lipset's *Political Man* (Baltimore: Johns Hopkins Press, 1981)—in equilibrium, democratic institutions must be both fair (all relevant political forces must perceive that they have a chance to win power in the future) and effective (they must make even losing under democracy better than a future under nondemocratic alternatives). Yet in some circumstances (for instance, those in which the power of the military is overweening), no such self-enforcing equilibrium can be achieved. And even under conditions that permit the establishment of fair and effective democratic institutions, other, less optimal, equilibria are also feasible. There is no guarantee that democracy will be a unique equilibrium even under the most favorable of circumstances, Przeworski concludes.

Chapter 2 discusses conditions under which liberalization—"an opening that results in the broadening of the social base of the regime without changing its structure" (p. 66)—occurs within an authoritarian regime.

On the basis of reasoning too complex to recount here, Przeworski concludes that liberalization is infeasible except under the heroic, but unrealistic, assumption that all the relevant political forces have full and accurate knowledge about everyone else's preferences and the probability that repression will be successful (p. 66). In the much more likely event that misperception occurs, either repression or full-scale transition to democracy is likely to result. The outcome in this case depends on a host of intangibles, including the preferences of the actors, the means by which these preferences are revealed, and their estimates of the efficacy of repression—all of which are likely to change once the process gets underway.

An intriguing comparison of the merits of capitalism and socialism is presented in chapter 3. The two kinds of systems are compared in terms of their normative appeal as social ideologies (or "blueprints"), and their feasibility of implementation. Capitalism offers a feasible blueprint, but one that is irrational because it cannot efficiently use and allocate all of its productive resources (under capitalism, some people are wont to starve). Socialism's blueprint is normatively more attractive to Przeworski because of its distributional implications, but it is simply impossible to put into practice (p. 122). Since both (pure) capitalism and (pure) socialism are found wanting, Przeworski favors a mix of the systems, but he can find no grounds on which to choose between social democracy and market socialism.

The final substantive chapter concerns the political consequences of economic reforms. Przeworski argues that economic reforms can advance far under democratic institutions, but that inevitably they will be politically destabilizing.

After digesting the implications of these various analyses, Przeworski bravely makes the pessimistic prediction that the East will become the South. "Political developments in Eastern Europe will not be different than in those countries where the transition to democracy occurred earlier, and economic transformations will stop far short of the blueprints" (p. 190). The only salient difference between these arenas is the East's geographic proximity to Western Europe. However, Przeworski doubts that this difference in geographic location is sufficient to overcome the shared deficits of weak state organizations, overregulated economies, inefficient agricultures, overweening public bureaucracies, rudimentary welfare systems, and emaciated political parties. (Although he does not mention it, the persistence of regional underdevelopment within the borders of states like England and Italy offers further reason to question the impact of geography on development prospects in Eastern Europe.)

One wonders, though, about the analytic path that leads to this pessimistic conclusion. Przeworski's game-theoretic apparatus treats both the interests and the solidarity of collective actors as exogenous, but since he tells us nothing about their determinants how does the analysis apply in given empirical cases? What are the interests of the military in Poland,

for example, and how solidary is it? That questions of this sort are not idle is evident, for the attempted Soviet coup of 1991 fell apart precisely because of divisions within the military.

Further, Przeworski fails to consider the possibility that geopolitical considerations might lead to substantial capital infusions—infusions that would decisively alter the payoffs for the relevant players in his games. On at least two counts, there are reasons to believe that geopolitics will have a different effect in the East than in the South.

In the first place, Eastern Europe has a more powerful political constituency in this country and in Europe (with the exception of Spain) than Central or Latin America. It is conceivable, if not likely, that one day this might translate into a greater willingness to subsidize the East in some contemporary version of the Marshall Plan. In the second place, there is a vastly greater potential for ethnoregional conflict in the East than in the South. This kind of conflict is especially likely to kindle the interest of powerful states. The threat of ongoing political instability in Eastern Europe (including in the present-day republics of the Soviet Union) may convince leaders of the G7 states to invest substantial amounts of capital there so as to guarantee a modicum of political stability. After all, when push came to shove, the Bush Administration preferred to live with Saddam Hussein than face the prospect of prolonged internal religioethnic conflict in Iraq. Are the geopolitical stakes in Eastern Europe and the USSR any less great than those in Iraq?

Manaus: Social Life and Work in Brazil's Free Trade Zone. By Leo A. Despres. Albany: State University of New York Press, 1991. Pp. xix + 322. $17.95.

Bryan R. Roberts
University of Texas at Austin

This is a fascinating study of the contemporary development of one of the world's most incongruous cities. Situated in the heart of the Amazon, Manaus first prospered at the turn of the century as the entrepot for the rubber boom, boasting an opera house and an electric tram system, but located at several weeks' distance by river travel from the nearest town. Languishing after the end of the boom, Manaus's urban decay illustrated the dangers of dependence on the vagaries of the world economy. Its current revitalization, based on export-oriented industrialization, is equally artificial. It has fewer economic and social linkages with its region than in the past and has a concentration of a much higher proportion of the area's population than it did when rubber tapping, trade, and transport were the bases of a centrifugal pattern of urbanization. As a free trade zone, benefiting from tax concessions, it has attracted modern assemblage industries that import almost all their inputs and export their

products and a large modern airport to cope with the thousands of shoppers who come from all over Brazil to buy the duty-free goods.

Despres carefully outlines the economic and political trends, both national and international, that have produced this urban anomaly, contributing an important case study of the role of cities in producing and reproducing underdevelopment. The study's major contribution is, however, at the microlevel. Despres provides data on samples of workers and their households chosen to illustrate the range of job situations in Manaus—electronics assemblage industry workers, wood-processing industry workers, white-collar workers in the service industries, and the self-employed. He analyzes income levels, household structure, consumption, and patterns of social interaction to provide an overview of lifestyles in the city.

Two major themes emerge—the considerable social heterogeneity of the city and the lack of any social order beyond the household and its immediate kinship relationships. About a third of households are extended, but this, argues Despres, is less a coping strategy, whether of survival or social mobility, than an eclectic response to various problems imposed by this particular urban context, that of migrants seeking shelter with kin, the scarcity of housing, the difficulties in finding work in a volatile labor market, and the desire to obtain an education. Even the extended family does not extend far, usually contains only close kin, and, like nuclear households, is not connected with wider networks.

The most common coping strategy is having multiple earners within a single household. Manaus is a situation of generalized poverty in which the majority of urban workers earn close to a minimum salary. Escaping the poverty trap in Manaus is done with difficulty and only in small households with multiple earners. These are the basis for an emerging middle class in which the parent(s) are in higher paid jobs and the children in white-collar occupations.

For most households, extra earners cannot compensate adequately for larger families, since the extra income is a marginal addition. Whether discussing the poor or the better-off, Despres shows the close connection between household life cycle and pattern of earnings. Males who are heads of households are likely to earn more than their wives and children, even when these are better-educated, and are more often employed in formal enterprises. The assemblage industries recruit heavily from young, unmarried dependents and have a high rotation of workers with little internal mobility. White-collar jobs, especially for women, offer incomes that are no better than those for manual work despite the workers' higher levels of education. Well-paid jobs in government or the private-service sectors tend to be filled by "spiralists," for whom working in Manaus is only part of a career that will take them to other parts of Brazil. Then there is the "informal" sector of the self-employed in which workers have considerable autonomy, prefer their status to that of wage employment, and earn, on average, more than those in the

assemblage industries and white-collar workers in commerce—who are often their children.

The city is also spatially heterogeneous. Slums invade residential districts, and there are few class-segregated housing estates. In this situation, Despres argues, neither industrial nor spatial organization creates the basis for community organization. Nor is there much evidence of institutions such as the church, either through the parish or through ecclesiastically based communities, filling this gap. Despres documents how Manaus's population is marginalized socially and politically by these processes, but what matters most to his informants is their exclusion from a reasonable standard of consumption. This is an excellent study, informing a wide variety of theoretical issues and providing a grounded view of some of the key population implications, particularly for household organization, of the current phase of dependent development.

Women, Sexuality and the Changing Social Order: The Impact of Government Policies on Reproductive Behavior in Kenya. By Beth Maina Ahlberg. Philadelphia: Gordon and Breach, 1991. Pp. xiv + 274.

Nadra Franklin and Antonio McDaniel
University of Pennsylvania

The diminishing influence of Kenyan women's groups on women's sexuality and fertility behavior is the subject of this interesting new book by Beth Maina Ahlberg. In Kenya, women's groups traditionally provided collective support in the regulation of sexuality and fertility. This historical source of power was dramatically transformed by the introduction of ideas about modern individualism. Today women's groups in Kenya primarily address household improvements and the workplace (pt. 1). This transformation signaled a loss of women's control in the area of fertility. Colonialism and "modernization" changed Kenya's traditional society by introducing a different, more Western, dimension to the marginalization and subordination of African women (pt. 2). Ahlberg illustrates these changes by closely examining case studies from her field research (pt. 3). Ahlberg concludes that the traditional collective spirit of cooperatives ensured that women's concerns about reproduction were accounted for, and the modern individualist model of reproduction has transferred fertility control to men.

Ahlberg speaks clearly and eloquently of the travesty of "modernization" in a voice that, historically, has been silenced. In her writing, she has provided affirmation for the validity and necessity of African women commenting on their own society. She lays claim to this right, confronting and exposing prevailing Western notions constructed about African women and their sexuality. Modernization in Kenya has actually meant that men gained control of their spouses' sexuality, although the general

assumption had been that, in traditional African society, the control of fertility decisions was in the hands of men. However, her results do confirm the view that nonindividualistic factors traditionally played a major role in fertility behavior in Africa.

The author aptly demonstrates her ability to combine historical research with the impassioned perspective of one who has witnessed the declining power of women within society. She has combined qualitative and quantitative survey methods to underscore the process in which women lose control of their sexuality. Ahlberg frames her work within the actor-systems dynamics model that looks upon the subordinating and constraining environment in which women's groups must function.

Ahlberg, a Kikuyu, conducted her field research among the Kikuyu women of Kenya, and used her status as an insider to her advantage. Knowledge of the language and local customs allowed Ahlberg to interpret the nonverbal clues during group discussion and to avoid being culturally suspect. In this, she stands in stark contrast to most Western anthropologists and sociologists who, claiming academic objectivity, presume to write about African culture from the perspective of an insider.

Integrating the demographic information collected in the survey would have strengthened the force of her argument, perhaps by showing that increasing Western influence resulted in increasing loss of women's control. Her study area is densely populated and was historically under the influence of colonial powers. This combination may have increased the ease with which domination occurred and influenced the response of this population to transformation more strongly. She provides data on the sex ratio that favors women, yet does not reflect on its possible interaction with women's loss of control in sexuality and fertility. Differences with respect to "modernizing" influences and the sex ratio between the study area and other areas in Kenya suggest caution in generalizing these particular conclusions to other areas of Kenya.

She selected study groups on the basis of the groups' meeting a distribution criterion that would include a variety of groups in the sample. Since this was not based on their proportionate representation in the study population but was done to provide variety, the sample does not accurately reflect the study area in its composition. As such, the force of any group's arguments, and the differences between the groups, should be put into the perspective of their proportionate representation in the larger population. In addition to case studies, Ahlberg surveyed the chairs of 177 groups registered with the government. However, we know nothing of the relative distribution of nonregistered groups compared with registered groups, nor if they both face similar constraints in collectives' activities. Also, differences between leaders who, because of government directives, are more highly educated and other group members may not be accounted for in the analysis.

The specificity of her data to Kikuyu women does not, however, detract from the force of her general conclusion that as a result of colonization and processes of westernization, African women lost control of their

own sexuality. There is something so very fundamental being said here that it runs the risk of being overlooked, and that is that the failure of family planning and health programs in Africa is due in large part to the general ignorance of and lack of respect for the local contexts and realities of Africa. Ahlberg gives us an excellent view of African women's loss of control over their sexuality and fertility from their own perspective and in their own context.

Japan's 'International Youth': The Emergence of a New Class of School-children. By Roger Goodman. Oxford: Clarendon Press, 1990. Pp. xv + 283. $65.00.

William K. Cummings
Harvard University

"Internationalization" is a central albeit murky term in recent Japanese discussions on educational reform. Roger Goodman provides valuable insight on these discussions through his focus on the policy changes devised to accommodate the youth returnees from overseas sojourns necessitated by their parent's posting to a foreign worksite. Goodman seeks to show how opinion in Japan has been socially constructed to magnify the supposed difficulties faced by these returnees, thus gaining them special privileges that include special returnee courses in public schools and exceptions to the stringent entrance requirements of leading universities. These privileges have been sought by the powerful business establishment in order to provide incentives for motivating dynamic businessmen to take up overseas assignments in the expanding Japanese economic empire.

Through the sixties, Japan's major overseas involvement was trade promotion of home-produced products and much of the international contact was carried out either from Japan or on short-term business trips. But from the mid-seventies the numbers of Japanese going to live overseas for business rapidly expanded. Some went to establish and manage new overseas factories, others to manage overseas financial institutions.

Their children often accompanied the high-status parents, though presumably this was at considerable risk, for once out of the Japanese educational hothouse, these young people might suffer the potential loss of their Japanese language abilities, not to mention their moral character (or such was the nature of the emerging myth). To prevent these losses, the Japanese government and overseas corporate groups set up various schooling options. In the Third World, the typical arrangement was a full-scale Japanese school covering all subjects and continuing through at least the lower secondary level—such schools served to isolate Japanese children from the negative characteristics of the native environment. In contrast, in developed societies, Japanese children were often encouraged

to attend local schools, while devoting Saturday to a Japanese school to keep up their language skills. Goodman finds that most children enjoyed their overseas sojourn, particularly the more relaxed school environment.

But coming home was another matter. Many experienced difficulty readjusting to Japanese life. Fellow schoolmates made fun of their foreign clothes and habits. Children were under considerable pressure to readjust. How severe were these difficulties? One of the most fascinating themes in Goodman's analysis is the wide gap between what the returnee children actually experienced and what the media and leading scholars said they experienced.

To put it simply, Goodman finds that the adjustment was relatively smooth for the majority of children, and that indeed most were advantaged by their overseas sojourn. They gained insight into foreign ways, became more expressive and social, and often gained reasonable competence in a foreign language—in sum, were "internationalized." Of course, the extent of these gains varied from child to child and rarely reached the ideal of a total cosmopolitan transformation.

While most returnees were not highly internationalized, neither did they experience tremendous adjustment problems. Their newly found expressive skills and their foreign-language skills enabled them to do well in their schoolwork, and most were even back to par in their Japanese within a year or so. One study finds that nonreturnees reported more "bullying" than did returnees—Goodman suggests that the overseas experience of the returnees led them to develop greater tolerance for rude antisocial behavior, so that they developed powers of adjustment up and above those of their homebound classmates. However, the media and many leading scholars persisted in stressing the difficulties of the adjustment process, of fortifying the "pitiful child" myth. Following from this myth were the various policies to ease the adjustment and to take advantage of the positive international qualities the returnees were assumed to possess.

Starting with this refreshing view on the returnee experience, Goodman then pursues several intriguing parallel analyses. For example, he raises the "comparative" question of why other problem groups such as the burakumin or Korean minorities are not given equally sympathetic treatment by the Japanese media. And why are not other returnee groups such as the children of Japanese who once immigrated to Brazil or the United States similarly accommodated? Each of these analyses lead him back to the central contention: the media looks after those who come from approved classes.

Perhaps the only parallel analysis I question is his case study in chapter 5 of a school that accommodates a number of returnee children. The school is clearly elite and private and different in many particulars from the schools where I have spent time. Yet Goodman asserts it is representative. Still much of the information he gains from this somewhat exceptional school does provide valuable insights on the adjustment experiences of the returnees.

Overall, Roger Goodman has demonstrated great creativity in his analysis of the returnee children phenomenon. He has convincingly demonstrated distortions in the factual reporting of this phenomenon by such major institutions as the scholarly research community, the mass media, and government commissions. And he has provided a plausible explanation for the class basis of these distortions. The analysis is well documented and written in a pleasing style, both scholarly and lively. One can learn much about the way modern Japan addresses a social problem from this excellent account.

The Color of Strangers, the Color of Friends: The Play of Ethnicity in School and Community. By Alan Peshkin. Chicago: University of Chicago Press, 1991. Pp. xiii + 304. $44.00 (cloth); $15.95 (paper).

David A. Kinney
University of Chicago

What sociohistorical situations and interpersonal processes allow adolescents to transcend their ethnic differences and come together as friends? *The Color of Strangers, the Color of Friends* is an engaging effort to address this question. Having previously studied the school-community relationship in two homogeneous settings, Alan Peshkin, a professor of education at the University of Illinois, Urbana-Champaign, chose a California city ("Riverview," population about 44,000) characterized by ethnic diversity initially to investigate which groups would be served by its single high school ("RHS"). In the process of finding that teens from the most economically disadvantaged backgrounds in the community (blacks and Hispanics) experienced the disproportionate share of academic problems, Peshkin uncovered success in RHS's social arena; a high degree of "naturalness and ease of cross-group contact—what students refer to as 'mingling'" (p. 190). Drawing on rich interview, observational, historical, and demographic data collected during the 1985–86 school year and from several one and two week visits in 1983–84 and 1984–85, Peshkin systematically investigated the nature and extent of "mingling."

Relevant demographic and historical information indicates that Riverview and RHS possess the relatively unique characteristic of not being dominated by any one ethnic group (high school student population: "white, 33 percent; black, 33 percent; Hispanic, 20 percent; Filipino, 12 percent; and the rest American Indian, Vietnamese, and others" [p. 290]). An important historical trend centers on the high level of unity and pride that has developed among Riverview residents in the face of being stigmatized as dangerous, poor, and black by outsiders living in wealthier, neighboring cities. Moreover, the turmoil and ethnic assertiveness that characterized both Riverview and the nation during the 1960s and 1970s gave way to harmonious interethnic associations in city

and school politics in the 1980s. Teachers were found to respect students' ethnicity, but tended to adopt "color-blind" stances in the classroom and focus on preparing students for adult work roles. Parents also view the school as an assimilating institution, while expressing preference for in-group marriages for their children.

Adolescents' search for identity included stops in both ethnic and nonethnic groups (e.g., "jocks, preppies, stoners"). The teens downplayed their ethnicity and, echoing the educators, commonly asserted, "People are people" (e.g., pp. 197–98). This philosophy was revealed in their close friendships and romantic relationships with peers from different ethnic groups. In addition to finding frequent interpersonal interactions across ethnic lines, Peshkin also found common "ethnic borrowing in regard to dialect, music, dance, and the like [and the] ordinary phenomenon of wannabes [who] transcend borrowing by their overall adoption of the behavior of another ethnic group, so that they act as if they were someone else" (p. 219). Overall, while these findings provide strong support for the author's argument for frequent mingling at RHS, a more extensive examination of the "wannabes" could possibly have led to some other interesting insights into contemporary adolescent socialization.

Peshkin also deals with the issue of blacks and Hispanics being underrepresented in the school's top academic track. He speculates that the school is not blamed for their lack of achievement because adolescents and adults "believe that black and Mexican students do less well in school compared to whites and Filipinos because their parents are less supportive of schooling. By so believing, the schools are not blamed for their failure. Second, black students are generally content with the school because they have a dominating presence [in the school's social scene] that lifts them into the school's limelight" (p. 253). Even though this issue is a secondary theme in the book, these patterns could have been discussed more fully by comparing them with John Ogbu's research on factors contributing to underachievement of minority students.

It is important to note that the convincing arguments put forth by Peshkin (see discussion of "begetting effects" [p. 255]) are further bolstered by his meticulous monitoring and explication of the role that his subjectivity played in this project (see the appendix, pp. 285–95 and pp. 4–5). In this refreshingly open and detailed account he chronicles the feelings that were evoked by his immersion in the field and how his values affected the research. Specifically, he delineates how his subjectivity both facilitated and precluded collecting different kinds of information. In addition, Peshkin provides suggestions for observing oneself so that researchers can "avoid the trap of perceiving just what [their] own untamed sentiments have sought out and served up as data" (p. 294).

This appendix should be considered by researchers from different disciplines and would be a provocative and instructive reading for students in methods courses. Overall, despite the minor limitations noted above, this is an important, thoughtful, and well-crafted book and is highly

recommended to sociologists of education, as well as to social scientists interested in community, race relations, and adolescent development.

The Limits of Law: The Public Regulation of Private Pollution. By Peter Cleary Yeager. New York: Cambridge University Press, 1991. Pp. xiii + 369. $49.50.

Lettie McSpadden Wenner
Northern Illinois University

The Limits of Law provides a well-researched, concise history of the evolution of attempts to reduce industrial pollution of U.S. waterways from 1948 through the 1980s. Peter Yeager presents a strong argument about the difficulties of obtaining social regulation because of the need that the state has to align itself with business interests in order to sustain its tax base. Scholars of law and society, he argues, should be as concerned with the formulation of laws as they are with the implementation of them because of the immense influence that industry has in the creation of statute law. He then develops a thorough analysis of the evolution of the Clean Water Act from its tentative beginnings of federal assistance for state regulation in the 1950s through its high point in 1972 when major amendments gave the U.S. Environmental Protection Agency (EPA) significant enforcement powers. However, he is less thorough in his coverage of the development of 1977 "mid-course corrections" in the Clean Water law and the 1987 amendments to this legislation. These two events, which set back deadlines and modified standards, certainly illustrate the power of industry to influence legislation more completely than the incremental increase in federal regulation from 1948 to 1972.

Most important, Yeager provides a thorough discussion of attempts to implement the Clean Water Act, where he shows that, even in the heyday of environmental enthusiasm, enforcement was less than complete. He takes us through the tortuous process by which industry objected to every step for achieving the best practicable treatment by the original 1977 deadline established by Congress. The steps were setting effluent standards, issuing specific permits and variances from permits to factories, monitoring the effluents, and enforcement proceedings. Repeated court challenges to each step resulted in delays well into the 1980s before even the first step, effluent standards for all industries, was in place. Since large companies had disproportionate legal and technical resources, their defensive tactics disadvantaged small firms, which proved to be easier targets for the regulators. Systematic discussion of enforcement efforts ends abruptly in 1978 when EPA ceased to collect and distribute data about administrative orders, notices of violations, and referrals to the Justice Department.

By emphasizing the lack of enforcement in the Carter years, the author

amply demonstrates his main point, that law per se is oriented to the status quo. Yet he clearly does not mean for us to abandon use of the courts to advance the cause of pollution control because the book contains several examples of how case law can advance that cause. One occurred prior to the passage of the 1972 amendments when citizen suits forced the creation of a permit system under the 1899 Rivers and Harbors Act; another, when a consent decree held EPA to its commitment to set standards for toxic pollutants beyond best practicable treatment. In the 1980s lack of governmental enforcement drove private groups to adopt the private attorney general role written into the 1972 act, which enables them to sue industrial polluters directly and is one of the remaining means of enforcement today.

Yeager has made a forceful argument that our political culture prevents government intrusion into proprietary processes, thus eliminating the one real option that could give regulators the power to change industrial processes in order to reduce the creation of pollutants. Instead they are restricted to add-on techniques for treating pollutants. The refusal by most administrations to fund governmental research on and development of innovative methods intensifies government's dependence on industrial knowhow for available solutions. He does not quite subscribe to the economists' solution of charging fees for effluents that are below standard to induce more innovative methods of preventing, rather than treating, pollution, but he offers us no alternative method of achieving this end. Yet he argues that the Clean Air Act's reliance on ambient air quality standards rather than on effluent standards may allow for more flexibility in solutions. That process, however, is even more convoluted than the one for water pollution control, adding stages to define what pollution is and to formulate state implementation plans, steps that afford industry even more opportunities to influence the outcome.

The deadline (1985) for there being no discharge into any navigable waterway is long past now, and it does not appear that the establishment of best practicable treatment has moved us any closer to best available or recycling standards. Rather our commitment to the add-on technology devised by industry may have preempted any further move to cleaner effluents, given the investment these sunk costs represent now. Yeager's clear description of the process by which a worthy goal can be derailed should be read by any student of the water pollution control laws.

Forging the Military-Industrial Complex: World War II's Battle of the Potomac. By Gregory Hooks. Urbana: University of Illinois Press, 1991. Pp. x + 303. $36.95.

G. William Domhoff
University of California, Santa Cruz

The misinterpretations, mistakes, and omissions are many in Gregory Hooks's attempt to support a "state-centered" theory of power in America through a study of the rise of an allegedly "autarkic" Pentagon due to industrial mobilization during World War II. The autonomy of the Pentagon is said to be based on its "control over the sources of investment capital and the letting of procurement contracts" (p. 5). This allows the Pentagon to be "a self-sufficient autarky operating independently of most other federal agencies" (p. 6). The defense program is said to be "first and foremost an exercise in industrial planning, with decision making lodged in the Pentagon" (p. 7). As for the professional military officers themselves: "the Pentagon's budgetary resources, power, and insulation, coupled with the distinctive worldview among professional military officers, allowed them to operate as 'relatively autonomous bureaucrats'" (p. 17).

Hooks sees the original basis for the Pentagon's autonomy in the strengthening of the state during the New Deal: "Administrative advances and strategic policy tools forged by the New Deal were absorbed by and adapted to the nascent national security state during and after World War II" (p. 5). He devotes 11 pages to the Department of Agriculture as a case of increasing state autonomy during the New Deal, which then declined primarily because state leaders turned their attention to the war agencies.

Although I disagree with almost everything Hooks says concerning the New Deal, and especially about the alleged autonomy of the Department of Agriculture from Southern plantation owners and the Farm Bureau, his first major mistake in terms of military autonomy is to equate the Pentagon with military control. Historians conclude that the "military bureaucracy" was run at the top by Wall Street lawyers and bankers. Many of those in the procurement bureaucracy under General Lucius Clay (son of a Georgia senator) were businessmen made instant officers. A Wall Street banker ran the Army-Navy Munitions Board at its key juncture and helped create the plan for allocating scarce resources to corporations. Hooks himself reports that Navy planners came from the "business community" (p. 116). At one point Hooks downplays the role of businessmen in uniform by saying they "served under career officers" (p. 148), but does not see that this completely undercuts his dismissal of the fact that these career officers were serving under Wall Street lawyers and bankers.

Hooks claims that General Brehon Somervell, head of the Army Service Forces, "resented" his civilian boss and "gained control" over

Army procurement (p. 96), but that is not what his citation actually says. Moreover, most sources report that Somervell fought New Dealers in the Planning Committee of the civilian War Production Board, not his civilian superiors. The real battles were between most businessmen and military men, on the one side, and New Dealers on the other, with few businessmen in the New Deal camp. The reorganization of the War Department that put Somervell in charge of the new Army Service Forces was carried out for the most part by management consultants, academic experts, and corporate lawyers.

Hooks's second big mistake is the misinterpretation of key decisions concerning war production. Barton Bernstein, Jordan Schwarz, and other historians show that businessmen dominated the mobilization. Businessmen early set their terms for participation; delayed defense production until America was at war; insisted on predominantly military production once the war began (even though New Dealers thought there could be more civilian production); stalled the gradual reconversion called for by New Dealers in 1943; and quickly removed controls as the war ended. The author nonetheless claims that military officers were the key actors in the major decisions. Part of the mistake is in the assumption that the "civilian" agency, the War Production Board, lost to the Pentagon, but it is wrong to equate business with the War Production Board and the military with the Pentagon (pp. 104–5). Businessmen were dominant at both places, although the most visible combatants in some arguments were New Deal economists at the War Production Board and military officers at the Pentagon. Business liked decision making at the Pentagon to freeze out New Dealers.

Hooks's third major mistake is to claim that there now exists a military-industrial complex separate from the "monopoly sector." He believes there is "little evidence" that firms specializing in postwar defense production are "closely interlocked with monopoly-sector firms" (p. 228), but there is excellent evidence that they are, including that in Stephen Johnson's study of the 15 largest defense firms ("How the West Was Won," *Insurgent Sociologist,* vol 6).

Hooks says he found the views of Mills "instructive," but in fact he reached his main conclusions without any use of Mills, as the bibliography of his 1985 dissertation reveals. He misinterprets Mills as a classical elitist too concerned with psychology, individuals, and the circulation of elites to put proper emphasis on institutions, but he and Mills actually have fairly similar views. The author also professes regard for the "corporate liberal" theory supposedly exemplified for World War II in the excellent research of Bernstein, who used both his interviews and archival data to show business ascendancy on several key issues. Hooks says this theory "must be rejected" because the mobilization was "set in motion according to strategic goals" (p. 93), not narrow business goals, as if anyone ever thought otherwise. The theory is caricatured with the insult that its analyses of the postwar period "tended" to see the defense program as a mere "subsidy" for large corporations (p. 34). It is taxed

with the claim that government financing of aircraft production "runs counter" to its "expectations" (p. 131); Hooks then shows that bankers were reluctant to finance businesses with no postwar civilian market, and that aircraft companies preferred government financing, which hardly counters any theory.

Hooks creates false issues when he contrasts "elites" and "structures," or assumes that business leaders only have profit motives when in government at the time of a major threat to the international capitalist system and democratic nation-states. His version of organizational theory narrows him to an economistic view of businessmen and a bureaucratic view of "state managers." Social power in general and class power in particular are lost from view. Like others focused exclusively on organizational power, Hooks is blind to class power and does not understand the interaction of the two. This book strikes me as empirically wrong and theoretically weak.

Representing Order: Crime, Law, and Justice in the News Media. By Richard V. Ericson, Patricia M. Baranek, and Janet B. L. Chan. Toronto: University of Toronto Press, 1991. Pp. xii + 383. $50.00 (cloth); $19.95 (paper).

Robert M. Entman
Northwestern University

Representing Order explores the coverage of law and crime in six Canadian media outlets. The book is, in important respects, innovative and provocative and full of useful observations on media content. Despite a noteworthy failure to connect its extensive and detailed content-analytical findings to larger theoretical concerns, the study makes a valuable contribution to the media studies literature.

The framing notion of the book is that the media represent social order through their coverage of the legal process and thereby reinforce public acceptance of ordering institutions and processes. The authors argue perceptively that media and legal institutions share important traits and social functions: "intertextually related to each other as part of a continuing dialogue about the terms and conditions of social order" (p. 342), the two institutions present constructed realities as natural and real, rely on fictional assumptions to reinforce norms that preserve order, and otherwise work together as social control mechanisms.

Yet the authors seek to go beyond the well-established notion that the media tend to support the status quo. They understand that the different media operate under different incentive structures and constraints and that these predict significant distinctions in their content. The book offers a framework that emphasizes the ways contrasting formats and markets yield varying news products. The chief formats are television, newspaper, and radio; the chief markets are "popular" and "quality." The six

media outlets that are studied include a popular and a quality outlet in each of the three formats.

Particularly praiseworthy here are the authors' trenchant critique of standard quantitative content-analytical technique, their insistence on combining quantitative and qualitative methods, and, most innovatively, their open acknowledgment of instability in their quantitative measures. Standard content analysis attempts to demonstrate its reliability, and hence its validity, by having news texts coded at least twice and then demonstrating an acceptable level of intercoder reliability. This often involves laborious training sessions in which coders are taught to see texts that they naturally read somewhat differently in the same way, so that the analysis will reflect the average response of a typical reader. The authors of this study argue that the variety of readings is itself theoretically relevant, and demonstrate this by including data indicating what dimensions of content showed statistically significant variation in readings among coders (most, in fact, did not). In its candor and creativity, its turning of methodological limitations into pertinent findings, this study deserves emulation.

A second, especially useful, feature of the book is its systematic attempt to assess differences in different media's coverage of the same news. This is one of the very few studies, and certainly the most thorough one, to take intermedia differences seriously. The attention to detail yields numerous specific insights, too many to include here, so I shall cite two examples. The data demonstrate that TV does not differ from print news in "presenting sources as normative witnesses to events occurring elsewhere and previously" (p. 350). Contrary to its self-presentation as "eyewitness" news, TV, like print, usually fails to cover events themselves but rather serves to transmit the views of certified observers about the events. As another example, the authors find that sources rarely provide evidence for their claims, beyond their own credibility; in this way the authors capture one of the markers of news rhetoric, one of the traits that distinguishes the persuasive techniques of mass media news from other, perhaps more rational and synoptic, forms of reasoning.

The authors share a large number of findings of this sort, about the media generally and an even larger number about differences among the media. It is in this respect that the book's weaknesses become apparent. There is simply too much detail in its content analysis; too little is digested, integrated, and linked to the promising theoretical agenda with which the book begins.

Simply put, the authors barely attend to the "So what?" question. The book hardly raises, let alone answers, the question of what difference it might make to society, the social order, or the legal process, that "quality" newspapers cover crime differently from "popular" ones. What are the social implications of distributing systematically different information about law and crime to the less-educated audience of popular media and the better-educated audience of quality media? Why do lower-status persons exhibit the tastes and create the market pressures

that induce the popular media to differ as they do from the quality? Here the work of Bourdieu could have been a model; indeed, the empirical data here are substantially richer than those Bourdieu has generally worked with, the issues raised quite similar, yet he is not even cited.

Rather than continue with these complaints, I would reiterate the book's particular merits: it is methodologically innovative and it probes intermedia differences with care that is much overdue. Such an ambitious project deserved an editor who might have persuaded the authors to condense their data presentation and focus it on the important theoretical issues. Attentive (and patient) readers will nonetheless find that the book's ability to raise important questions repays the effort of reading it.

The Journalism of Outrage: Investigative Reporting and Agenda Building in America. By David L. Protess, Fay Lomax Cook, Jack C. Doppelt, James S. Ettema, Margaret T. Gordon, Donna R. Leff, and Peter Miller. New York: Guilford, 1991. Pp. xvii + 301. $30.00.

Stephen Hilgartner
Columbia University

The Journalism of Outrage is a study of investigative reporting and its role in shaping policy agendas in the United States. The seven authors, most of whom are professors of journalism or communication, conducted six case studies of print or television exposés. In each, they obtained the cooperation of the news organization prior to the release of the story. The extent of access varied, but in several cases, field researchers directly observed the process of constructing the investigative report. Knowing about forthcoming stories also allowed the authors to conduct survey research (with a quasi-experimental, preexposure/postexposure design) to see whether stories had an impact on public opinion and on the opinions of elites.

By far, the most interesting feature of this book is the authors' examination of "what reporters do" to move an "embryonic investigative story" to publication (p. 204). Several of the case studies are splendid, and each reveals the twists and turns that shaped "the" story as it evolved. The authors examine how competition among media outlets influenced the selection of story ideas. They describe how reporters "pitched" stories to editors to win backing for expensive investigations. They explore conflict within news teams about the proper "angle" of a story and about the "genre" of wrongdoing. They document the hunt for compelling "villains" and sympathetic "victims," and show how finding individuals to cast in these roles can play a decisive role in determining how a story ends up. They discuss the search for documentation that will make charges stick. They explore the reporters' interactions with expert sources, who often try to influence the focus of exposés.

And they provide hilarious examples of the extremes to which television reporters will go to get exciting visuals. Toward the end of the book, the authors make some generalizations about the "life course" of exposés. A rich picture emerges of how contingencies and organizational and dramaturgical factors determine what kinds of investigative stories are likely to survive in this arena.

Unfortunately, the theoretical analysis of the influence of investigative reports in building policy agendas is disappointing. The authors expend most of their energy attacking what they call the "mobilization model" of how investigative reporting influences policy. This three-box, two-arrow model holds that "published exposés outrage the general public, who in turn demand and often get reforms from government officials" (p. xii). The authors emphasize that the model is linear, unidirectional, and sequential (first the exposé outrages the public, then the public demands change, then officials respond to pressure). Another key feature of the model is that the demands for reform emerge from the *general* public, not from networks of policy operatives.

As if to justify the space they spend attacking the mobilization model, the authors claim that it is prevalent in the journalism community. Perhaps. The authors support their contention by pointing out that this model fits neatly with conventional views of American democracy and by providing some illustrative quotations from the world of journalism. (The example that most closely mirrors their own representation of the mobilization model comes from a 1989 advertisement, run by the Society of Professional Journalists, extolling the virtues of a free press.) Even so, one wonders whether most journalists actually hold such a schematic, linear view of how reforms are enacted. For journalists, this way of representing the influence of exposés may operate more as a succinct, literary gloss, or as a useful myth, than as a literal description of the process.

Social science narratives often use a good foil to sharpen their focus and intensify drama. Every argument needs an opponent, but the mobilization model is too weak to be cast as a credible villain. It treats "the public" as an undifferentiated mass, lacking social organization; casts government officials in a passive, reactive role; makes no allowance for the activities of interest groups or social-movement organizations; neglects the conclusions of constructivist research on social problems; and ignores organizational network analyses of policy domains. Because the authors concentrate on combating the mobilization model, they miss many opportunities to deploy their data in theoretical contests that are less one-sided. More interesting debates—in fields ranging from political sociology, to social problems theory, to the sociology of culture and knowledge—are ignored.

By the end of the book the authors have scored many points against the mobilization model, and they conclude that "policy reforms result from complex interactive factors rather than from the simple linear progression" that the model describes (p. 250). For sociologists, this finding

is hardly "news." Nevertheless, the empirical data and analysis of the life course of investigative reports remain very interesting. Students of mass media, culture, social problems, and public policy will find this material intriguing and useful as they pursue their own research agendas.

Architecture: The Story of Practice. By Dana Cuff. Cambridge, Mass.: MIT Press, 1991. Pp. xi+306. $24.95.

Michael Saenz
University of Texas at Austin

In this, the first book "to report extensive empirical data from observations within architectural practice" (p. 15), Dana Cuff offers an intimate account of the profession's everyday institutions and ethos. Drawing preeminently on Geertz, she attempts a thick description of architecture's "culture" by organizing her observations around "rich stories . . . that will ring true to insiders," and then examining those stories for the meaning they hold for interested parties (p. 10). She succeeds in documenting that—professional ideology aside—"the design of our built environment emerges from" unwieldy, open-ended, highly negotiated "collective action" (p. 12).

Cuff's observations were made during six months of fieldwork with three southern California firms, and are supplemented by 50 additional interviews and her experience as an architecture student, teacher, and consultant. Her stories form the core of chapters describing "design problems in practice, . . . the metamorphosis of layperson into architect" (p. 15), and the social milieu of the architectural office and its clients. She also offers a more evaluative chapter tracing the history of three projects regarded as excellent by the profession, clients, and public alike.

Each discussion treats what Cuff regards as "dialectical dualities" structuring architects' self-conceptions. She suggests that professional ideology one-sidedly emphasizes the individual over the collective, art over business, and design as individual decision making over design as collaborative sense making. As a result, Cuff suggests, architects commonly ignore or struggle vainly against inevitable components of their actual conditions of practice. Cuff usually prescribes solutions for the profession's dilemmas; while many of her recommendations are not new, they gain new force from the care and detail of her descriptive studies.

Consider, for example, the crucial activity of design. Cuff identifies six dimensions of real-world design, all but one of which go unspoken or denied by architectural training and professional dogma. The admitted characteristic is "design in the balance": in practice, the physical form of a building is the architect's "primary professional value," to be defended in every negotiation. But design in practice is also influenced by "countless voices"—the renderer choosing a view for a presentation

drawing, a compromising contractor, or the bank officer hesitant about financing. Practical design evolves too within a climate of "professional uncertainty," in which the responsibilities, procedures, and claims to expertise revolving around the production of something so complex as a building remain ultimately ambiguous. Design in practice equally involves a process of "perpetual discovery," in which the information required to make final decisions is never certain, complete, or agreed upon. Actual design is further subject to "surprise endings," because in the ongoing relationship between client, architect, and others, "the object of the negotiations, the building, is created within the negotiations" (p. 95)—a particularly difficult admission for professional ideology, which regards design as the masterful, rational, individual elaboration of a formal design premise. And, since buildings entail considerable expenditures and people live with them for a long time, design in practice is a "matter of consequence," raising emotional as well as financial stakes, and inviting Machiavellian intrigue. Cuff suggests that if architects are to practice their trade effectively, the profession will have to reconstrue its formalist notion of design to include the practice's consensual, processual, and contingent aspects. This means extending full professional recognition to architects who manage, negotiate, and finance, as well as to those who labor over the drawing board, and incorporating such topics in architectural education.

This sensitive discussion, and others like it, are likely to come as welcome revelation to anyone accustomed to approaching design as the aesthetic mystery (and architecture as the artistic calling) presented in academic studios, art history courses, and architectural magazines. Indeed, Cuff's observations both evoke the nature of architectural practice and recast it in a way that makes it freshly scrutable—the twin requisites for successful interpretative ethnography.

Cuff keeps her cultural analysis admirably focused on identifiable rituals, specifiable professional beliefs, and concrete activities. The theoretically inclined reader might expect a fuller discussion of practice than Cuff's brief allusion to Geertz, or the one-line operational definitions she offers ("routine activities based on commonplace experience"). Some consideration of Bourdieu's ideas of practice might have been particularly apt, since his notions of habitus and the social production of distinction would lend themselves well to explicating architects' highly studied differentiation by school of design, firm, and individual life-style.

Extended theoretical discussion, however, is not the intent of Cuff's book. She is more concerned with characterizing ideal-typical conditions of professional practice. Her second chapter, moreover, presents a short, excellent social history of the profession (incorporating a laudable review of the literature) to frame her contemporary observations. Overall, her efforts result in a remarkable analysis of the implicit assumptions and manifest beliefs governing architectural practice—a long-overdue articulation of the profession's unspoken but widely enacted contradictions.

*The Rude Hand of Innovation: Religion and Social Order in Albany,
New York, 1652–1836.* By David G. Hackett. New York: Oxford Uni-
versity Press, 1991. Pp. xv + 240. $29.95.

Rhys H. Williams
Southern Illinois University, Carbondale

David Hackett's book is a largely successful attempt to wed social history
with current culture theory. He pursues the relationship between religion
and social order by exploring the changes in both in Albany, New York.
Religion influences social order differently in diverse historical periods,
Hackett argues. He illustrates this point by contrasting Albany's colonial
period with its life during the revolutionary and early national periods.

During the colonial period, Albany's social life was dominated by its
Dutch-based Calvinism. The Dutch church stood, literally and meta-
phorically, in the center of the town and was responsible for the oversight
and welfare of the entire community: "[the] sacred bond between elders
and congregation was similar to the civil bond between Common Council
members and citizens" (p. 20). Further, religion offered only a limited
range of possibilities for meanings and interactions and thus both cultural
and social life were relatively stable. Newcomers, while usually not
Dutch, were incorporated into the community through Dutch social
networks and religious values. Family networks were of particular
importance in ensuring both commercial connections and community
membership.

There were tensions between local Dutch culture and English colonial
authority. Rather than consider themselves part of the war effort against
the French, Albany's town fathers (literally) assumed their primary re-
sponsibility was to the physical and economic well-being of "their" com-
munity; British troops were accepted only to the extent they could be
assimilated into local culture and communal ties. Dutch cultural practices
were retained within English legal forms with the town's leadership play-
ing the role of mediators. The inherent tension was controlled to the
extent to which the community itself was culturally integrated.

But the fissures became more overt. Increased English immigration
threatened the coherence of Dutch control, and ethnic antagonisms in-
creased. Resentment of the British Crown was also increasing, however,
with the Crown's actions being interpreted as incursions into communally
held privileges, rather than violations of individual liberties. Internal
ethnic antagonisms gave way to a local alliance against "external" En-
glish impositions.

The threat of British invasion "shifted the basis for local political unity
from Dutch ethnicity to American citizenship" (p. 39) and became the
local origins of American nationalism. Political circumstances forged the
townfolk into a common identity as Christians and Americans rather
than as Dutch or Calvinists. Hackett is particularly persuasive in tracing

the political, economic, and ideological context for this transformation that well illustrates the changing role of religion in a changing society.

The political crises became the impetus for new interpretations of community membership and a new role for the churches (plural). Responsibility for communal welfare shifted from the church to the town as the political sphere was broadened and differentiated. But the unity was short-lived. By the 1790s Albany experienced a "Yankee invasion" and increased conflict over ethnicity, inequality, morality, and the meaning of community. The "rude hand of innovation" upended a cultural system long dominated by tradition.

Concomitantly, religion's role in social order also changed. Religions became "ideologies," explicit meaning systems that had a direct and immediate impact on social behavior. Dutch Calvinism had to compete with Presbyterianism, Methodism, and emerging secular ideologies such as republicanism. Albany's culture became a contested arena and religion changed from the implicit framework that organized all communal life to an ideological resource with which various social groups presented themselves to the community. It is an interesting story, well told here. I was left, however, with a few questions.

Hackett challenges "modernization" theory's bipolar movement from gemeinschaft forms of community to gesellschaft forms of association. This approach particularly overestimates the linearity of religion's decline or marginalization. And the richness of Albany's history provides ample evidence of resurgence and innovation; religion remained an important source of cultural meanings and social solidarity, even when the content of both changed.

But while that is no doubt true, it is also true that modernization theory is not a very difficult target for the social historian committed to "local knowledge." The more difficult task is offering an alternative. Hackett prefers the cultural analyses of Geertz and Ann Swidler. While I am sympathetic to this perspective I have questions about the explicit distinction between the way culture "works" in settled periods of stability versus its role in unsettled periods of change and contention. In the former, religion is "culture," an implicit framework influencing behavior indirectly and apparently noncognitively. In the latter, religion becomes an "ideology," influencing behavior "directly" and used more or less cognitively as a political resource.

Thus, rather than arguing that historical specificity constantly alters religion's role in meaning creation and interpretation, the implication is that there are two types of historical periods, settled and unsettled, and two attendant modes of influence. We are back to a bipolar social environment—but, instead of a linear transition from traditional to modern, we now have oscillation between settled and unsettled. Why should this distinction produce these differing modes, and how do we establish whether any given period is settled or unsettled? Is the distinction justified through theoretical necessity (i.e., the nature of culture) or empirical

induction? Certainly the approach improves upon modernization theory, but not as clearly as is promised.

Of course, these questions should not be laid at Hackett's feet alone. As it is, he has given us a rich social history that effectively engages some of the most important questions in the current study of religion, political culture, and social theory. Historians and sociologists alike will benefit from reading his efforts.

You Are My Darling Zita. By Glenn Busch. Philadelphia: Temple University Press, 1991. Pp. 304. $29.95.

Steve Gold
Whittier College

A growing body of work in sociology, anthropology, and literary criticism has recently questioned the stance of traditional ethnographic writing, characterizing it as excessively academic, impersonal, and autocratic. A major accusation has been that in both style and content, ethnographies are more concerned with justifying their writer's authority—his or her wit, insight, and believability—than in dealing with subjects' experiences in a detailed and straightforward manner. The result, according to this line of criticism, is that traditional ethnographies often wind up reinforcing the same nonreflexive, positivistic views of the world and power inequalities between researcher and subject that the ethnographic tradition is supposed to avoid.

Critics of the traditional ethnographic approach have called for the creation of "a new ethnography." They argue that scholars must take the risk of using innovative and experimental techniques to develop a means of writing ethnographies that will avoid the problems that plague this style of research.

Glenn Busch is not a sociologist and does not mention the new ethnography in the introduction to *You Are My Darling Zita.* Nevertheless, the book is an example of what might be considered an experimental ethnography, and a successful one at that. The book consists of in-depth oral histories of six elderly New Zealanders, who reflect on their lives. Adding richness, Busch supplements his transcripts with photographs of interviewees taken at various times during their lives. Some of these images have been lifted from family albums, while others were taken by the author.

Eschewing any mention of sociological theory and not offering analysis of the interviews, the author lets his subjects speak for themselves. In addition to his avoidance of the ethnographic convention of analyzing his interviews, Busch also rejects the traditional introduction of a sociology text—one that emphasizes current social issues and reviews relevant literature. Instead, he describes his motives for writing *You Are My Darling*

Zita as being rooted in his own personal life—his confronting the loss of his grandmother by reestablishing links with elderly folk.

In relying on oral histories collected from people representing various walks of life, and in its exclusion of the interviewer, this book is analogous to Studs Terkel's classic *Working* (New York: Pantheon Press, 1974) except that it deals with the elderly instead of workers. The interviews, collected from ordinary New Zealanders (many of whom were born in the British Isles), are rich and varied. Respondents are middle class and working class, men and women, single and married, old fashioned and "new age," and have and do not have children. Their occupations range from coal miner to social worker. The last interview represents a married couple who offer distinct perspectives on their shared life.

In different ways, each respondent describes his or her growing up in and then dealing with an incredibly rigid and unforgiving, affection-free Victorian culture that might have been taken, whole cloth, from a Dickens novel. The subjects suffered from cruel stepparents and foster parents, mine disasters, repressed sexuality, and childhood illnesses. One man was forever banished from learning his father's trade and forced into a life in the mines because he left his workbench one afternoon to play ball. Despite the harshness of their lives, respondents describe moments of happiness as well—in environs, friendship, family, religion, and love.

The book's subjects reflect on life, love, religion, sex, and their own mortality with insight and candor. Because most are infirm, their views on the loss of physical ability and the dependency on caregivers that results are especially informative.

This book is far from the conventions of mainstream sociology and is written by a nonsociologist. Nevertheless, it is eminently readable and poignant. The photographs included with each oral history add to the book's depth. After hearing about these people's relations with their parents, you really can see these relationships represented in family snapshots. However, the photographs would have contributed more if the interviewees had been asked to reflect on them, when and where they were taken, who is shown along with the protagonist, and how they felt about the "self" that is represented in the photograph.

You Are My Darling Zita would serve well in a course on aging, sociology of the life course, or Anglo-Saxon culture to enrich and humanize the concepts offered in these fields. It offers scholars and students alike powerful evidence that valuable knowledge about society can be gained simply by talking and listening to the people around them.

Looking Awry: An Introduction to Jacques Lacan through Popular Culture. By Slavoj Zizek. Cambridge, Mass.: MIT Press, 1991. Pp. ix + 188. $22.50.

Edith Kurzweil
Rutgers University

Slavoj Zizek is gifted and versatile: he was a researcher at the Institute of Sociology in Ljubjana, and had run as a proreform candidate for the presidency of the republic of Slovenia, before writing *Looking Awry*—a book he "conceived as a kind of introduction to Lacanian 'dogmatics,' . . . as an excuse for indulging in the idiotic enjoyment of popular culture" (pp. vii–viii). He took time off from his political commitments in order to indulge in "post-deconstructionism"—his strictly Lacanian reading of the "Real" (in relation to the "Symbolic" and the "Imaginary").

Zizek assumes that his readers are conversant with all of Lacan's works, that they are as taken up with Lacanian psychoanalysis and its roots in linguistic theory as he is himself. This means, also, that the child is primarily conceptualized as both by having been born into its language and by the ambivalence deriving from the dual (and coexisting and inseparable) nature of signs and signifiers, and that, therefore, its apprehension of the world derives from all the circumstances and impressions surrounding language learning. In discussing, for instance, "The Hitchcockian Blot as the Gaze of the Other," Zizek refers to "the Lacanian notion of the master signifier . . . that does not denote some positive property of the object but establishes, by means of its own act of enunciation, a new intersubjective relation between speaker and hearer" (p. 103). In fact, all of his film criticism is grounded in the assumption that his readers will be tantalized by such neo-Lacanian associations and fabrications.

Thus Zizek assumes we know that the elusive *"objet petit a"* refers to the inferior size of the little boy's penis which, in the Imaginary of the child extrapolates to his feelings of inferiority to his father; is intrinsic to all of his future emotions and perceptions of conscious and unconscious reality; and of patriarchal society—which is dominated by fathers who possess the phallus, determine laws of inheritance and the relations between families and partners, and so forth. (Such a simple explanation of "libidinal economy," of course, "misunderstands" Lacan but may aid the uninitiated who read this short review.)

At one point, Zizek locates Sam Spade, the hero of Dashiell Hammett's *Maltese Falcon,* in "the function of the *objet petit a* at its purest, . . . [in] that elusive make-believe that drove the man to change his existence" (p. 8); at another point, he moves via Hegel and "Freud's obsession with Michelangelo's Moses" and "simple causes [that] can produce 'chaotic' behavior," to theories of intuition in physics, especially of the "strange attractor." Here, this attractor turns into "a kind of physical metaphor

for the Lacanian *objet petit a* . . . [that] draw[s] us into chaotic oscillation"—into the antithesis of order (p. 38).

Zizek's book is neither sociology, anthropology, or any other social science; it is not a story or narrative; nor is it fiction or psychoanalysis, or what used to be called criticism: it belongs to postmodern theory—itself a slippery term that even its proponents have problems defining. But insofar as postmodernism is described as full of ruptures, breaks, and discontinuities, *Looking Awry* exemplifies it: Zizek moves elegantly from film to film, from text to text. For instance, he explains that, in Hitchcock's *The Trouble with Harry,* the social life of the village goes on as people exchange pleasantries while viewing Harry's corpse:

> Just like the obsessive personality described by Freud toward the end of his analysis of the "Rat Man," so the "official ego" of the characters in *The Trouble with Harry,* open, tolerant, conceals a network of rules and inhibitions that block all pleasure . . [in a body that] is present without being dead on the symbolic level . . [so that] the only denoument the story can have is Harry's symbolic death. [Pp 26–27]

And we learn that "Harry's problem is the same as Hamlet's—the drama of a real death unaccompanied by a symbolic 'settling of accounts'."

Clearly, Zizek also supposes that we are familiar with the immense Lacaniana that has been growing astronomically since Lacan's death in 1982, especially in literary theory and in the new field of film criticism. *All* of Zizek's analyses of the productions of popular culture are Lacanian, with a dash of Foucault (particularly in the use of the "gaze"). The reader alternates between admiration and dizziness as he or she moves from one ambiguous utterance to another, from hidden meanings and connections to double entendres, identifications, and misidentifications in, among others, Hitchcock's films *Rear Window, North by Northwest, Secret Agent, Sabotage,* and so on; and as he follows the juxtapositions of Hitchcock's artistic aims and multiple texts, his plots and protagonists' gestures, their fantasies and innuendos, and comparisons with films by Fritz Lang, James Cameron, Sergei Eisenstein, and a score of other filmmakers who surpass each other in their abilities to build up suspense.

Among "postmodernists" who judge this type of literary fantasy in terms of imaginative leaps by practitioners, Zizek is bound to be admired: he reaffirms their own activities as he moves, for instance, from *Hamlet* to *Antigone,* and maintains that "understatement becomes a specific way of taking note of the 'blot' created by the real of the paternal body" (p. 27).

Like Lacan, Zizek easily moves from the human condition to language and to the various ways our symbolic edifice is so easily shattered— sometimes erupting in spectacular form, as in the radiation emanating from Chernobyl, which "confronted us with the threat of what Lacan calls 'the second death': the result of the reign of the discourse of science

is that what was at the time of the Marquis de Sade a literary fantasy (a radical destruction that interrupts the life process) has become today a menace threatening our everyday life" (p. 36).

But Zizek keeps his promise: he does exemplify Lacan's theories and in the process provides us with excellent summaries of film plots, with many parallels to Lacan's analysis of Poe's "Purloined Letter," and he elucidates the types of *jouissance* we may experience when watching a thriller. As required by postmodernists, he makes his bows to radical politics without engaging in political action. Inevitably, Zizek's intellectual footwork seems somewhat farfetched when, for example, he parallels Marx's concept of *surplus value* with *surplus enjoyment* and then to exchanges between capital and labor. I kept thinking: Why does not Zizek continue to use his energies in the Slovenes' struggle? For the language of postmodernism is already outmoded, while his society certainly needs the help of one of its most intelligent and talented sons.

Hollywood Shot by Shot: Alcoholism in American Cinema. By Norman K. Denzin. New York: Aldine de Gruyter, 1991. Pp. xvii + 292. $39.95 (cloth); $22.95 (paper).

Alexander Hicks
Emory University

In this ambitious, multifaceted book Norman Denzin interprets the Hollywood "alcoholism film" as a genre, as a reflection of the "cultural structure" and "lived experience" underlying alcoholism, and as a constituent of this same culture and experience. Within a unifying framework that highlights reciprocal relations between film and society, the study is theoretically eclectic. The delineation of the alcoholism genre is structuralist in the manner of Will Wright's *Sixguns and Society* (Berkeley and Los Angeles: University of California Press, 1976). Analysis of the genre's social origins employs reflection theory. Both Denzin's delineation of the genre and his interpretations of its films draw on historical approaches to social problems and deviance. In addition, Denzin's readings of reviewers' "hegemonic" interpretations of the films draw on reception theory. Denzin rolls high. He plays for a landmark work and runs the risk of inelegance.

In chapter 1, Denzin provides a preliminary sketch of the sequenced elements that define the alcoholism genre, a synopsis of his framework, a preview of major interpretive themes and conclusions ("drink, but not to excess"), and a description of his purposive sample of films. This consists of 40 films judged representative of the development of the genre through six periods and phases: temperance (1909–20), prohibition (1921–34), preclassic (1935–45), classic (1946–59), interregnum (1960–80), and present (1980–). In chapter 2, Denzin analyzes Hollywood's

durable but, for Denzin's genre, tangential treatment of the diverting "happy alcoholic" from Chaplin through *Harvey* and *Arthur 2*.

In part 2, Denzin enters into detailed discussion of the preclassic and classic phases of his genre. Hollywood's search for the "alcoholic hero" and its less manifestly heroic depiction of the female alcoholic are deciphered and compared by drawing on the *Star Is Born* cycle (1932, 1937, and 1954) and the celebrated *Lost Weekend* for heroes (in chap. 3) and on *Smash-Up* and *I'll Cry Tomorrow* for heroines (in chap. 4). Although redemption from alcoholism means "hitting the bottom" for all, the female alcoholic falls lower and rises back up less autonomously. As the physician/psychiatrist supplants the "good spouse" as savior, the male Oedipal complex fades away but the female complex merely gains one more patriarch. In chapter 5, *Come Back Little Sheba* and several other films are used to document the 1950s preliminary recognition of the familial dynamics of and the Alcoholics Anonymous (AA)/institutional perspective on alcoholism.

After his excursion in chapter 6 through the 1962–80 disintegration of antialcoholic censure, of the alcohol/drug distinction, and of the unalloyed alcoholism genre during the 1962–80 "interregnum," Denzin returns to his main business with a reconsideration of alcoholic family, heroine, and hero in the reconsolidated, if transformed, genre of the 1980s. As chapter 7 relates, the family now ranges from succor (*Tender Mercies*) to victim (*fatal* victim in *Under the Influence*); and the AA/medical "disease" definition of alcoholism triumphs along with AA/medical control of the diseased. As chapter 8 documents, the female now attains new independence in the struggle against alcohol, as in *The Morning After,* but the male is now customarily granted major *second,* nonalcoholic identities such as lawyer (*The Verdict*), policeman (*8 Million Ways to Die*), athlete (*Fat City*), or the like. The new period is entropic: while *Barfly* dares to celebrate the experience of unreformed alcoholics, *Clean and Sober* and *My Name Is Bill W.* return us to the classical cycle of grace, fall, and redemption.

In his concluding chapter, Denzin structurally models the "classic" alcoholism film in terms of a detailed sequence of elements. He fails to model nonclassical phases of the genre with equal tidiness, but he does rather sharply summarize the overall genre. Denzin's attempts to provide subversive feminist readings of the genre in terms of the "male gaze" and "Oedipal journey" are too sketchy and seem rushed. However, he introduces a seminal reinterpretation of the "alcoholism film" as a subset of the "love story" in which the drinking problem provides the narrative tension and triumph over drink provides a romantic resolution cleansed of sex. Denzin never provides the tidy match of historical period and genre phase that seems promised early on, and he delivers little on his promise to explore effects of the alcoholism film on the lived experience of alcoholics. However, his historical interpretation of reviewers' responses to alcoholism films and alcoholism subtly reads the times, yield-

ing such revelations as reviewers' long-delayed abandonment of a "weakness of will" explanation for the disease.

As he risked with his swollen agenda, Denzin, by biting off more than he can chew, does indeed fall victim to inelegance. However, Denzin's overreaching by no means prevents his providing hardy nourishment for sociologists of culture, deviance, social problems, and film. Indeed, Denzin's ambitions, eluded as well as realized, should inspire further cultural treatments of deviance and social problems. They might also revive a moribund—not to impugn the current profusion of loosely sociological "cultural studies"—a comatose sociology of film. For this neglected sociological genre Denzin provides the best work since Will Wright's 1976 classic and a work of some sure influence and durability in its own right.

ACKNOWLEDGMENTS TO REFEREES

Keith Hope
Alexandra Hrycak
Sheri Hseuh
Joan Huber
Jay Hughes
David Huizinga
Albert Hunter
Larry Isaac
Marilyn Ivy
David Jacobs
Jerry A. Jacobs
Leif Jensen
Jason B Jimerson
Hank Johnston
Arne L. Kalleberg
Mark G. Kelman
David Kerbow
Lynne Kiesling
Yong-Hak Kim
Michael S. Kimmel
Edgar Kiser
Evelyn Kitagawa
Bert Klandermans
Alden S. Klovdahl
Fred Kniss
David Knoke
Melvin Kohn
Claire H. Kohrman
George Kourvetaris
David Kowalewski
Marvin D Krohn
David Laitin
Nancy Landale
John Laub
Shin Ying Lee
Eric M. Leifer
Donald Light
Nan Lin
James R Lincoln
Bruce G Link
Colin Loftin
Helena Z Lopata
Diane Lye
Stewart Macaulay
David Maines
Robert D. Mare
Margaret Mooney
   Marini
Barry Markovsky
Peter Marsden
Ross L. Matsueda

Armand L. Mauss
Susan E. Mayer
John McAdams
John McCarthy
Meredith McGuire
Sara McLanahan
Hugh B. Mehan, Jr.
Scott Menard
Steven F. Messner
John W. Meyer
Ruth Milkman
Beth A. Mintz
John Mirowsky
Phyllis Moen
Harvey Molotch
James Montgomery
S. Philip Morgan
Martina Morris
Edward Murguia
John Myles
Robert L. Nelson
Charles Noble
William M. O'Barr
Pamela Oliver
Susan Olzak
Suzanne T. Ortega
J Scott Parrott
Misagh Parsa
Orlando Patterson
Silvia Pedraza
Charles Perrow
Trond Petersen
Richard A. Peterson
Jeffrey Pfeffer
Frances Fox Piven
Alejandro Portes
Dudley L. Poston
Daniel Powers
Jill Quadagno
Charles Ragin
Katherine Ragsdale
James Ranger-Moore
Craig Reinarman
Michael J. Reynolds
Peter J. Richerson
Cecilia Ridgeway
Martin Riesebrodt
Matilda Riley
Ronald R. Rindfuss
Everett Rogers
Rachel A. Rosenfeld

Alice Rossi
Guenther Roth
David Rowe
Gilbert F. Rozman
Dietrich
   Rueschemeyer
Steve Rytina
Nader Saiedi
Arthur Sakamoto
Jimy M. Sanders
Andrew A.
   Sappington
Masamichi Sasaki
Saskia Sassen
Joachim J. Savelsberg
Derek Sayer
Emanuel Schegloff
Carmi Schooler
James T. Schulman
Pepper Schwartz
David Sciulli
W. Richard Scott
Gerald W. Scully
William H. Sewell,
   Jr.
Louise Shelley
Alan Sica
Bernard S. Silberman
Susan S. Silbey
Brian D. Silver
Theda Skocpol
Wesley Skogan
John Skvoretz
Shelley A. Smith
C Matthew Snipp
Aage Sørensen
Glenna Spitze
Susan Sprecher
Mark C Stafford
Charlotte Steeh
Suzanne K. Steinmetz
John D. Stephens
Haya Stier
C. Shannon Stokes
Anselm Strauss
Sheldon Stryker
Michael S. Swafford
Richard Swedberg
Ivan Szelenyi
Szonja Szelenyi
Richard P. Taub

CONTENTS OF VOLUME 97

Articles

Commentary and Debate

Review Essays

Book Reviews

1776 ERICSON, RICHARD V., PATRICIA M. BARANEK, JANET B. L. CHAN. *Representing Order: Crime, Law, and Justice in the News Media.* Robert M. Entman

226 ETZIONI-HALEVY, EVA. *Fragile Democracy: The Use and Abuse of Power in Western Societies.* Betty A. Dobratz

1481 EVERS, A., and H. WINTERSBERGER, eds. *Shifts in the Welfare Mix: Their Impact on Work, Social Services and Welfare Policies.* Jerald Hage

1156 FAURÉ, CHRISTINE. *Democracy without Women: Feminism and the Rise of Liberal Individualism in France.* Joan W. Scott

571 FINNEGAN, RUTH. *The Hidden Musicians: Music-Making in an English Town.* K. Peter Etzkorn

1194 FISKE, ALAN PAGE. *Structures of Social Life: The Four Elementary Forms of Human Relations.* John F. Stolte

549 FLIGSTEIN, NEIL. *The Transformation of Corporate Control.* Lauren B. Edelman

537 FOLEY, DOUGLAS E. *Learning Capitalist Culture: Deep in the Heart of Tejas.* Paula Brush

1502 FONT, MAURICIO A. *Coffee, Contention, and Change in the Making of Modern Brazil.* Richard Graham

1487 GALANTER, MARC, and THOMAS PALAY. *Tournament of Lawyers: The Transformation of the Big Law Firm.* Daniel Poor

1489 GARLAND, DAVID. *Punishment and Modern Society: A Study in Social Theory.* John R. Sutton

216 GELBER, STEVEN M., and MARTIN L. COOK. *Saving the Earth: The History of a Middle-Class Millenarian Movement.* Anson Shupe

1751 GERGEN, KENNETH J. *The Saturated Self: Dilemmas of Identity in Contemporary Life.* Philip Wexler

559 GHOSH, BAHNISIKHA. *The Indian Population Problem: A Household Economics Approach.* George Farkas

853 GIVEN, JAMES. *State and Society in Medieval Europe: Gwynedd and Languedoc under Outside Rule.* Bruce G. Carruthers

883 GLUCKSMANN, MIRIAM. *Women Assemble: Women Workers and the New Industries in Inter-War Britain.* Vicki Smith

1506 GOBALET, JEANNE G. *World Mortality Trends since 1870.* George J. Stolnitz

BOOK REVIEWERS

Princeton

Politics and Jobs

The Boundaries of Employment Policy in the United States

Margaret Weir

Americans claim a strong attachment to the work ethic and regularly profess support for government policies to promote employment. Why, then, have employment policies gained only a tenuous foothold in the United States? To answer this question, Margaret Weir highlights two related elements: the power of ideas in policymaking and the politics of interest formation. Rather than seeing policy as a straightforward outcome of public preferences, she shows how ideas frame problems and how interests form around possibilities created by the interplay of ideas and politics.

"*Politics and Jobs* establishes a new landmark in the study of economic and social policies in the United States. Weir's insightful and analytically powerful book will be widely read and cited for years."—*William Julius Wilson, University of Chicago*

Cloth: $29.95 ISBN 0-691-07853-X

Thorstein Veblen and His Critics, 1891-1963

Conservative, Liberal, and Radical Perspectives

Rick Tilman

The influential economist and philosopher Thorstein Veblen was one of the most original and penetrating critics of American culture and institutions, and his work still attracts the attention of scholars from a wide range of political viewpoints and scholarly disciplines. Focusing on the doctrinal and theoretical facets of Veblen's political economy, this book offers a study not only of his ideas but also of the way his critics have responded to them. As he scrutinizes the ideologies of the conservatives, liberals, and radicals who commented on Veblen, Rick Tilman portrays the diversity of social theory in the first half of the twentieth century.

Cloth: $39.50 ISBN 0-691-04286-1

Moving Crucifixes in Modern Spain

William A. Christian, Jr.

Why are religious visions believed only in certain times and places? In this book William Christian investigates the settings and responses to a series of group visions reported by Spaniards in rural Galicia, Valencia, Cantabria, and Navarre in the early part of this century. In light of the social strife and strong anticlerical movements of the period, the author examines how gender and religious politics influenced the experiences of seers and the interpretation of their visions by church officials, journalists, and the public.

Cloth: $24 95 ISBN 0-691-07387-2

New in paperback

The Chances of Explanation

Causal Explanation in the Social, Medical, and Physical Sciences

Paul Humphreys

This book provides a post-positivist theory of deterministic and probabilistic causality that supports both quantitative and qualitative explanations. Features of particular interest include the ability to provide true explanations in contexts where our knowledge is incomplete, a systematic interpretation of causal modeling techniques in the social sciences, and a direct realist view of causal relations that is compatible with a liberal empiricism.

"This work is by far the best thing written on probabilistic causality and probabilistic explanation to date."
—Wesley C. Salmon, University of Pittsburgh
Now in paper: $14.95 ISBN 0-691-02086-8

Princeton University Press

41 WILLIAM ST. • PRINCETON, NJ 08540
ORDERS: 800-777-4726 • OR FROM YOUR LOCAL BOOKSTORE

Human Studies is devoted primarily to advancing the dialogue between philosophy and the human sciences. In particular, such issues as the logic of inquiry, methodology, epistemology and foundational issues in the human sciences exemplified by original empirical, theoretical and philosophical investigations are addressed. Phenomenological perspectives, broadly defined, are a primary, though not an exclusive focus.
Human Studies is attractive to scholars in a variety of fields, since it provides a forum for those who address these issues in attempting to bridge the gap between philosophy and the human sciences. The contributions published have been drawn from sociology, psychology, anthropology, history, geography, linguistics, semiotics, ethnomethodology, political science and, of course, philosophy.
Human Studies is partially supported by the Department of Sociology, Boston University, Boston, Massachusetts.

Journal
Highlight

KLUWER
ACADEMIC
PUBLISHERS

AN INVITATION TO
REFLEXIVE SOCIOLOGY

Pierre Bourdieu and Loïc J. D. Wacquant

This book is the first and only comprehensive introduction to the work of Pierre Bourdieu, among the world's most influential and imaginative social theorists. Systematic and accessible, it is the ideal entree into the core of his wide-ranging *oeuvre*.

The structure of Bourdieu's theory of knowledge, practice, and society is first dissected by Loïc Wacquant; he then collaborates with Bourdieu in a dialogue in which they discuss central concepts of Bourdieu's work, confront the main objections and criticisms it has met, and outline Bourdieu's views of the relation of sociology to philosophy, economics, history, and politics. The final section captures Bourdieu in action in the seminar room as he addresses the topic of how to practice the craft of reflexive sociology.

Pierre Bourdieu Gasland/SIPA Press

Paper $13.95 (est.) 264 pages (est.)
Library cloth edition $38.95 (est.)

THE UNIVERSITY OF CHICAGO PRESS
5801 SOUTH ELLIS CHICAGO, IL 60637

CHICAGO

THE RADIANT PAST

Ideology and Reality in Hungary's Road to Capitalism

Michael Burawoy and János Lukács

"*The Radiant Past* is a stunning book destined to become a classic. It combines fascinating ethnographic research of the highest quality with theoretical subtlety and profound political relevance. The icing on the cake is that it is tremendously fun and engrossing to read, not just for specialists on Eastern Europe or industrial sociology, but for anyone who likes to read sociology at its best." — Erik Olin Wright, University of Wisconsin, Madison

Cloth $24.95 240 pages

CAPITALIST DEVELOPMENT AND DEMOCRACY

Dietrich Rueschemeyer, Evelyne Huber Stephens, and John D. Stephens

"This is one of the finest recent works of comparative sociology. The authors cover a vast canvas of the world's countries, systematically reviewing the stance of various social classes to democracy over the last one hundred years. . . . I can recommend the book most strongly to all those interested in the development of democracy — in the past and in the world today." — Michael Mann, University of California, Los Angeles

Paper $19.95 375 pages Cloth $45.00

MAYORS AND MONEY

Fiscal Policy in New York and Chicago

Ester Fuchs

Mayors and Money explores the political dimension of urban fiscal problems and reveals the importance of a mayor's role in a city's response to economic crisis.

Paper $15.95 248 pages Cloth $42.00

American Policy and Political Economy series

MAKE ROOM FOR TV

Television and the Family Ideal in Postwar America

Lynn Spigel

Spigel shows how the arrival of television, a massive movement of entertainment into the privacy of the home, absorbed and reflected deep cultural anxieties about family life in the postwar period.

Paper $15.95 288 pages Cloth $42.00

CROSSING THE POSTMODERN DIVIDE
Albert Borgmann

> "*Crossing the Postmodern Divide* is exhilarating in the
> sweep of its argument, the vigor of its language, and the
> vividness of its examples. . . . I think we are in the mood for
> an end of the century, end of millennium, reflection on
> what we have come to and where we want to go.
> Borgmann's book is an important contribution to this
> reflection." — Robert Bellah
> Cloth $19.95 184 pages

THROUGHOUT YOUR GENERATIONS FOREVER
Sacrifice, Religion, and Paternity
Nancy B. Jay

> An extraordinary feminist study of relations between sacrifice,
> gender, and social organization. "This capacious, erudite,
> wide-ranging book is worthy of its Durkheimian and Weberian
> lineage and will stand as one of the great sociologies of
> religion and great feminist theories." — Nancy J. Chodorow,
> University of California, Berkeley
> Cloth $27.50 224pages
> 2 halftones, 4 line drawings, 1 table

LANGUAGE LOYALTIES
A Source Book on the Official English Controversy
Edited and with an Introduction by James Crawford

> "A complete chronology extending from founding father
> Benjamin Franklin's views on the topic to those of S. I.
> Hayakawa. . . . This book is a *must* for anyone wishing
> to understand the historical and current perspectives
> surrounding an official English policy in the United States."
> — Amado M. Padilla, Stanford University
> Paper $14.95 536 pages Cloth $45.95

CHICAGO

INTENSIVE CARE
Medical Ethics and the Medical Profession
Robert Zussman

In riveting case studies, Zussman describes how medical
decisions in intensive care units are negotiated and made.
He concentrates on the practice of medical ethics—how
definitions of right and wrong emerge from the social situa-
tions of patients, families, doctors, and nurses and from the
workings of hospitals and the courts. From his superb field-
work—observing medical staff on their rounds; interviewing
staff, patients, and families; and systematically reviewing
hospital records—Zussman reveals the existence of deep
conflicts of opinion on how to allocate treatment and
resources.
Cloth $29.95 (est.) 256 pages (est.)
11 line drawings, 3 tables

GENERAL EDUCATION IN THE SOCIAL SCIENCES
Centennial Reflections on the College
of the University of Chicago
John J. MacAloon

Since its inception sixty years ago, Social Science 2 has been
one of the most significant courses in American undergraduate
education. Here, fourteen eminent social scientists examine
its intellectual excitement and reflect upon the larger issues
of the theory and practice of general education.
Paper $12.95 (est.) 296 pages (est.)
Library cloth edition $42.00

A Centennial Publication

Sommaire

--✂-----